Residuals and goodness-of-fit tests for stationary marked Gibbs point processes

Jean-François Coeurjolly, Frédéric Lavancier

Residuals and goodness-of-fit tests for stationary marked Gibbs point processes

Jean-François Coeurjolly[1,2] and Frédéric Lavancier[3]
[1] Gipsa-lab, CNRS, Grenoble, France
[2] Laboratory Jean Kuntzmann, Grenoble, France.
[3] Laboratoire Jean Leray, Nantes, France.

Abstract

The inspection of residuals is a fundamental step to investigate the quality of adjustment of a parametric model to data. For spatial point processes, the concept of residuals has been recently proposed by Baddeley et al. Baddeley et al. (2005) as an empirical counterpart of the *Campbell equilibrium* equation for marked Gibbs point processes. The present paper focuses on stationary marked Gibbs point processes and deals with asymptotic properties of residuals for such processes. In particular, the consistency and the asymptotic normality are obtained for a wide class of residuals including the classical ones (raw residuals, inverse residuals, Pearson residuals). Based on these asymptotic results, we define goodness-of-fit tests with Type-I error theoretically controlled. One of these tests constitutes an extension of the quadrat counting test widely used to test the null hypothesis of a homogeneous Poisson point process.

AMS 2000 subject classifications: Primary 62M30, 60G55; secondary 60K35, 62F03, 62F05, 62F12

Keywords: stationary marked Gibbs point processes, residuals, goodness-of-fit test, quadrat counting test, maximum pseudolikelihood estimator, Campbell Theorem, Georgii-Nguyen-Zessin formula, central limit theorem for spatial random fields

1 Introduction

Recent works on statistical methods for spatial point pattern makes parametric inference feasible for a wide range of models, see Møller (2008) for an overview of this topic and more generally the books of Daley and Vere-Jones (1988), Stoyan et al. (1987) Møller and Waagepetersen (2003) or Illian et al. (2008) for a survey on spatial point processes. The question is then to know whether the model is well-fitted to data or not. For classical parametric models, this is usually done via the inspection of residuals. They play a central role in parametric inference, see Atkinson (1985) for instance. This notion is quite complex for spatial point processes and has been recently proposed by Baddeley et al. Baddeley et al. (2005), following ideas from a previous work of Stoyan and Grabarnik Stoyan and Grabarnik (1991).

The definition of residuals for spatial point processes is a natural generalization of the well-known residuals for point processes in one-dimensional time, used in survival analysis (see Fleming and Harrington (1991) or Andersen et al. (1993) for an overview). For example, a simple measure of the adequacy of a one-dimensional point process model consists in computing the difference between the number of events in an interval $[0, t]$ and the conditional intensity (or hazard rate of the lifetime distribution) parametrically estimated and integrated from 0 to t. The extension in higher dimension requires further developments due to the lack of natural ordering. It may be done for point processes admitting a conditional density with respect to the Poisson process. These point processes correspond to the Gibbs measures. The equilibrium in one dimension between the number of events and the integrated hazard rate may be replaced in higher dimension by the *Campbell equilibrium* equation or *Georgii-Nguyen-Zessin* formula (see Georgii (1976), Nguyen and Zessin (1979a) and Section 2.3), which is the basis for defining the class of $h-$residuals where h represents a test function. In particular, Baddeley et al. Baddeley et al. (2005) consider different choices of h leading to the so-called raw residuals, inverse residuals and Pearson residuals, and show that they share similarities with the residuals obtained for generalized linear models.

Thanks to various diagnostic plots developped in the seminal paper Baddeley et al. (2005) and implementation within the `R` package `spatstat` Baddeley and Turner (2005), residuals appear to be a very convenient tool in practice. Some properties of the residuals process are exhibited in Baddeley et al. (2005) and Baddeley et al. (2008), including a conditional independence property and variance formulae in particular cases. In these two papers, the authors conjecture that a strong law of large numbers and a central limit theorem should hold for the residuals as the sampling window expands.

Our paper addresses this question for $d-$dimensional stationary marked Gibbs point processes. We obtain the strong consistency and the asymptotic normality in several contexts for a large class of test functions h. The $h-$residuals crucially depend on an estimate of the parameter vector. We consider the natural framework where the estimate is computed with the same data over which the $h-$residuals are assessed. The assumptions are very general and we show that they are fulfilled for several classical models, including the area interaction point process, the multi-Strauss marked point process, the Strauss type disc process, the Geyer's triplet point process, etc. The assumptions on the estimator are quite natural and we show that they are fulfilled in particular by the maximum pseudolikelihood estimator (in short MPLE) (see Baddeley and Turner (2000) for instance), for which asymptotic properties are now well-known (see Jensen and Møller (1991), Jensen and Künsch (1994), Billiot et al. (2008), Dereudre and Lavancier (2009) and Coeurjolly and Drouilhet (2009)).

Moreover, based on these asymptotic results, we propose statistical goodness-of-fit tests for which the Type-I error is asymptotically controlled. To the best of our knowledge, this is the first attempt in this direction. Such tests exist for rejecting the assumption of a homogeneous or inhomogeneous Poisson point process, but for general marked Gibbs point processes, the existing validation methods are either graphical (for example by using the QQ-plot proposed by Baddeley et al. (2005)) or rely on Monte-Carlo based simulations. We present two tests based on the computation of the residuals on different subdomains of the observation window. They extend in a very natural way the quadrat counting test for homogeneous Poisson distributions (see Diggle (2003) for instance). Besides, we present a test which combines several different $h-$residuals (associated to different functions h), computed on the entire observation window. The next step will be to implement these testing procedures to assess their power, compare them and reveal their limits. A thorough study will require extensive simulations and should deserve a separate paper.

The rest of the paper is organized as follows. Section 2 gathers the main notation used in this paper and briefly displays the general background. The definition of marked Gibbs point processes is given. They depend exclusively on the choice of an energy function or equivalently, of a local energy function. All the assumptions are based on this function. The *Georgii-Nguyen-Zessin* formula is recalled, leading to the definition of the $h-$innovations and $h-$residuals. Some examples are presented including the classical residuals considered by Baddeley et al. (2005) and new ones connected to the well-known empty space function (or spherical contact distribution), denoted in the literature by F, see Møller and Waagepetersen (2003) for instance.

Section 3 deals with asymptotic properties and presents our main results. A parametric stationary d-dimensional marked Gibbs point process is observed in a domain, denoted by Λ_n, assumed to increase up to $\mathbb{R}^d$. Sufficient conditions expressed in terms of the test function and the local energy function are given in order to derive the strong consistency result. We also propose an asymptotic control in probability of the departure of the h-residuals process from the h-innovations through the departure of the estimate from the true parameter vector (see Proposition 3). This allows us to deduce asymptotic normality results. Two different frameworks are considered: for the first one, the initial domain is splitted into a fixed finite number of subdomains (with volumes aimed at converging to $+\infty$) and we consider the vector composed of the $h-$residuals computed on each subdomain. For the second framework, we consider the vector composed of the h_j-residuals (for $j = 1, \dots, s$) computed on the same domain Λ_n, where $h_1, \dots, h_s$ are different test functions.

The asymptotic normality results depend on unknown asymptotic covariance matrices. The important question of estimating these matrices is addressed in Section 4. We give a general condition under which these matrices are definite-positive and propose a consistent estimate.

Section 5 exploits the asymptotic results obtained before. Some goodness-of-fit tests are proposed, based on normalized residuals computed in the two previous frameworks. They are shown to converge to some χ^2 distribution. Framework 1 leads to a generalization of the quadrat counting test for homogeneous Poisson distributions. Framework 2 yields a test which combines the information coming from several residuals, as for instance residuals coming from the estimation of the empty space function at several points.

The different assumptions made in the previous sections to obtain asymptotic results are discussed in Section 6. When considering classical test functions, exponential family models and the MPLE, the regularity and integrability type assumptions are shown to be satisfied for a wide class of examples. The testing procedures require moreover an identifiability condition to provide a proper normalization. Proposition 16 shows that this condition is easy to check for the first proposed test and appears to be not restrictive in this case. For the other tests, checking this condition depends more specifically on the model and the test function. We show, in Proposition 18, how this condition can be verified on two examples of models with several choices of test functions.

In Section 7, the very special situation where the energy function is not hereditary is considered. The GNZ formula is not valid any more in this setting but, provided a slight modification, it has been recently extended in Dereudre and Lavancier (2009). This leads to a natural generalization of the residuals to the non-hereditary setting.

Proofs of our main results are postponed to Section 8. The main material is composed of an ergodic theorem obtained by Nguyen and Zessin (1979b) and a new multivariate central limit theorem for spatial processes. Our setting actually involves some non stationary conditional centered random fields. A general central limit theorem adapted to this context has been obtained in Comets and Janzura (1998) for self-normalized sums (see also Jensen and Künsch (1994) in the stationary case and without self-normalization). But, contrary to these papers where the observation domain is assumed to be of the form $[-n, n]^d$, we consider domains that may increase continuously up to $\mathbb{R}^d$. This particularity, which seems more relevant, requires an extension of the results in Comets and Janzura (1998) and Jensen and Künsch (1994) to triangular arrays. This new central limit theorem is presented in Appendix A.

2 Background on marked Gibbs point processes and definition of residuals

2.1 General notation, configuration space

We denote by $\mathcal{B}(\mathbb{R}^d)$ the space of bounded Borel sets in $\mathbb{R}^d$. For any $\Lambda \in \mathcal{B}(\mathbb{R}^d)$, Λ^c denotes the complementary set of Λ inside $\mathbb{R}^d$. The norm $|.|$ will be used without ambiguity for different kind of objects. For a vector $\mathbf{x}$, $|\mathbf{x}|$ represents the uniform norm of $\mathbf{x}$; For a countable set $\mathcal{J}$, $|\mathcal{J}|$ represents the number of elements belonging to $\mathcal{J}$; For a set $\Delta \in \mathcal{B}(\mathbb{R}^d)$, $|\Delta|$ is the volume of Δ.

Let $\underline{\mathbf{M}}$ be a matrix, we denote by $\|\underline{\mathbf{M}}\|$ the Frobenius norm of $\underline{\mathbf{M}}$ defined by $\|\underline{\mathbf{M}}\|^2 = Tr(\underline{\mathbf{M}}^T\underline{\mathbf{M}})$, where Tr is the trace operator. For a vector $\mathbf{x}$, $\|\mathbf{x}\|$ is simply its euclidean norm.

For all $\mathbf{x} \in \mathbb{R}^d$ and $\rho > 0$, $\mathcal{B}(\mathbf{x}, \rho) := \{\mathbf{y},\ |\mathbf{y} - \mathbf{x}| < \rho\}$. Let us also consider the short notation, for $i \in \mathbb{Z}^d$, $\mathbb{B}_i(\rho) = \mathcal{B}(i, \rho) \cap \mathbb{Z}^d$.

The space $\mathbb{R}^d$ is endowed with the Borel σ-algebra and the Lebesgue measure λ. Let $\mathbb{M}$ be a measurable space, which aims at being the mark space, endowed with the σ-algebra $\mathcal{M}$ and the probability measure λ^{m}. The state space of the point processes will be $\mathbb{S} := \mathbb{R}^d \times \mathbb{M}$ measured by $\mu := \lambda \otimes \lambda^{\mathrm{m}}$. We shall denote for short $x^m = (x, m)$ an element of $\mathbb{S}$.

The space of point configurations will be denoted by $\Omega = \Omega(\mathbb{S})$. This is the set of simple integer-valued measures on $\mathbb{S}$. It is endowed with the σ-algebra $\mathcal{F}$ generated by the sets $\{\varphi \in \Omega,\ \varphi(\Lambda \times A) = n\}$ for all $n \in \mathbb{N}$, for all $A \in \mathcal{M}$ and for all $\Lambda \in \mathcal{B}(\mathbb{R}^d)$. For any $x^m \in \mathbb{S}$ and $\varphi \in \Omega$, we denote $x^m \in \varphi$ if $\varphi(x^m) > 0$. For any $\varphi \in \Omega$ and any $\Lambda \in \mathcal{B}(\mathbb{R}^d)$, we denote $\varphi_\Lambda := \varphi_{\Lambda \times \mathbb{M}}$ the projection of φ onto $\Lambda \times \mathbb{M}$, which is just the mesure $\sum_{x^m \in \varphi \cap (\Lambda \times \mathbb{M})} \delta_{x^m}$, where δ_x is the Dirac measure at x. We will use without ambiguity some set notation for elements in Ω, e.g. $\varphi \cup \{x^m\} = \varphi \cup x^m := \varphi + \delta_{x^m}$ and for $x^m \in \varphi$, $\varphi \setminus \{x^m\} = \varphi \setminus x^m := \varphi - \delta_{x^m}$. For any $\Lambda \in \mathcal{B}(\mathbb{R}^d)$, the number of elements of φ_Λ is denoted by $|\varphi_\Lambda| := \varphi(\Lambda \times \mathbb{M})$.

2.2 Marked Gibbs point processes

The framework of this paper is restricted to stationary marked Gibbs point processes. Since we are interested in asymptotic properties, we consider these point processes on the infinite volume $\mathbb{R}^d$. Let us briefly recall their definition.

A marked point process Φ is a Ω-valued random variable, with probability distribution P on $(\Omega, \mathcal{F})$. The most prominent marked point process is the marked Poisson process π^ν with intensity measure ν on $\mathbb{R}^d$ (and mark density λ^{m}). The homogeneous marked Poisson process arises when $\nu = z\lambda$, with $z > 0$.

Let $\theta \in \mathbf{\Theta}$, where $\mathbf{\Theta}$ is some compact set of $\mathbb{R}^p$ (for some $p \geq 1$). For any $\Lambda \in \mathcal{B}(\mathbb{R}^d)$, let us consider the parametric function $V_\Lambda(.;\theta)$ from Ω into $\mathbb{R} \cup \{+\infty\}$. For fixed θ, $(V_\Lambda(.;\theta))_{\Lambda \in \mathcal{B}(\mathbb{R}^d)}$ constitutes a compatible family of energies if, for every $\Lambda \subset \Lambda'$ in $\mathcal{B}(\mathbb{R}^d)$, there exists a measurable function $\psi_{\Lambda,\Lambda'}$ from Ω into $\mathbb{R} \cup \{+\infty\}$ such that

$$\forall \varphi \in \Omega \quad V_{\Lambda'}(\varphi;\theta) = V_\Lambda(\varphi;\theta) + \psi_{\Lambda,\Lambda'}(\varphi_{\Lambda^c};\theta). \tag{1}$$

From a physical point of view, $V_\Lambda(\varphi_\Lambda;\theta)$ is the energy of φ_Λ in Λ given the outside configuration φ_{Λ^c}. The following definition is the classical way to define Gibbs measures through their conditional specifications (see Preston (1976)).

Definition 1. *A probability measure P_θ on Ω is a marked Gibbs measure for the compatible family of energies $(V_\Lambda(.;\theta))_{\Lambda \in \mathcal{B}(\mathbb{R}^d)}$ and the intensity ν if for every $\Lambda \in \mathcal{B}(\mathbb{R}^d)$, for P_θ-almost every outside configuration φ_{Λ^c}, the law of P_θ given φ_{Λ^c} admits the following conditional density with respect to π^ν:*

$$f_\Lambda(\varphi_\Lambda|\varphi_{\Lambda^c};\theta) = \frac{1}{Z_\Lambda(\varphi_{\Lambda^c};\theta)} e^{-V_\Lambda(\varphi;\theta)},$$

where $Z_\Lambda(\varphi_{\Lambda^c};\theta)$ is a normalization called the partition function.

The existence of a Gibbs measure on Ω which satisfies these conditional specifications is a difficult issue. We do not want to open this discussion here and we will assume that the Gibbs measures we consider exist. We refer the interested reader to Ruelle (1969); Preston (1976); Bertin et al. (1999); Dereudre (2005); Dereudre et al. (2010), see also Section 6 for several examples.

In this article, we focus on stationary marked point processes on $\mathbb{S}$, i.e. on point processes admitting a conditional density with respect to the homogeneous marked Poisson process π. Moreover, without loss of generality, the intensity of the Poisson process, z, is fixed to 1. We assume in a first step that the family of energies is hereditary, which means that for any $\Lambda \in \mathcal{B}(\mathbb{R}^d)$, for any $\varphi \in \Omega$, and for all $x^m \in \Lambda \times \mathbb{M}$,

$$V_\Lambda(\varphi;\theta)) = +\infty \Rightarrow V_\Lambda(\varphi \cup \{x^m\};\theta)) = +\infty, \tag{2}$$

or equivalently, for all $x^m \in \varphi_\Lambda$, $f_\Lambda(\varphi_\Lambda|\varphi_{\Lambda^c};\theta) > 0 \Rightarrow f_\Lambda(\varphi_\Lambda \setminus \{x^m\}|\varphi_{\Lambda^c};\theta) > 0$. The non-hereditary case will be considered in Section 7. The main assumption is then the following.

> **[Mod-E]**: For any $\theta \in \mathbf{\Theta}$, the compatible family of energies $(V_\Lambda(.;\theta))_{\Lambda\in\mathcal{B}(\mathbb{R}^d)}$ is hereditary, invariant by translation, and such that an associated Gibbs measure P_θ exists and is stationary. Our data consist in the realization of a point process with Gibbs measure $P_{\theta^\star}$. The vector $\theta^\star$ is thus the true parameter to be estimated, assumed to be in $\mathring{\mathbf{\Theta}}$.

The local energy to insert a marked point x^m into the configuration φ is defined for any Λ containing x^m by

$$V\left(x^m|\varphi;\theta\right) := V_\Lambda(\varphi \cup \{x^m\}) - V_\Lambda(\varphi).$$

From the compatibility of the family of energies, i.e. (1), this definition does not depend on Λ. We restrict our study to finite-range interaction point processes, which is the main limitation of this paper.

> **[Mod-L]**: There exists $D \geq 0$ such that for all $(m,\varphi) \in \mathbb{M} \times \Omega$
>
> $$V\left(0^m|\varphi;\theta\right) = V\left(0^m|\varphi_{\mathcal{B}(0,D)};\theta\right).$$

2.3 Definitions of residuals for spatial point processes

The basic ingredient for the definition of residuals is the so-called GNZ formula stated below.

Theorem 1 (Georgii-Nguyen-Zessin Formula)**.** *Under **[Mod-E]**, for any function $h(\cdot,\cdot;\theta) : \$ \times \Omega \to \mathbb{R}$ (eventually depending on some parameter θ) such that the following quantities are finite, then*

$$\mathbf{E}\left(\int_{\mathbb{R}^d\times\mathbb{M}} h\left(x^m,\Phi;\theta\right) e^{-V(x^m|\Phi;\theta^\star)}\mu(dx^m)\right) = \mathbf{E}\left(\sum_{x^m\in\Phi} h\left(x^m,\Phi\setminus x^m;\theta\right)\right), \tag{3}$$

where $\mathbf{E}$ denotes the expectation with respect to $P_{\theta^\star}$.

For stationary marked Gibbs point processes, (3) reduces to

$$\mathbf{E}\left(h\left(0^M,\Phi;\theta\right) e^{-V\left(0^M|\Phi;\theta^\star\right)}\right) = \mathbf{E}\left(h\left(0^M,\Phi\setminus 0^M;\theta\right)\right) \tag{4}$$

where M denotes a random variable with probability distribution λ^{m}. The following definition is based on empirical versions of both terms appearing in (4).

Definition 2. *For any bounded domain Λ, let us define the $h-$innovations (denoted by I_Λ) and the $h-$residuals (denoted by R_Λ and depending on an estimate $\widehat{\theta}$ of $\theta^\star$) by*

$$\begin{aligned}
I_\Lambda\left(\varphi;h,\theta^\star\right) &:= \int_{\Lambda\times\mathbb{M}} h\left(x^m,\varphi;\theta^\star\right) e^{-V(x^m|\varphi;\theta^\star)}\mu(dx^m) - \sum_{x^m\in\varphi_\Lambda} h\left(x^m,\varphi\setminus x^m;\theta^\star\right) \\
R_\Lambda\left(\varphi;h,\widehat{\theta}\right) &:= \int_{\Lambda\times\mathbb{M}} h\left(x^m,\varphi;\widehat{\theta}\right) e^{-V\left(x^m|\varphi;\widehat{\theta}\right)}\mu(dx^m) - \sum_{x^m\in\varphi_\Lambda} h\left(x^m,\varphi\setminus x^m;\widehat{\theta}\right).
\end{aligned}$$

From a practical point of view, the last notion is the most interesting since it provides a computable measure. The main examples considered by Baddeley et al. in Baddeley et al. (2005) (in the context of stationary point processes) are obtained by setting $h(x^m, \varphi; \theta) = 1$ for the raw residuals, $h(x^m, \varphi; \theta) = e^{V(x^m|\varphi;\theta)}$ for the inverse residuals and $h(x^m, \varphi; \theta) = e^{V(x^m|\varphi;\theta)/2}$ for the Pearson residuals. In particular, one may note that the raw residuals constitutes a difference of two estimates of the intensity of the point process (up to a normalisation by $|\Lambda|$): the first one is a parametric one and depends on the model while the second one is a nonparametric one (since it is equal to $|\varphi_\Lambda|$). Another more evolved example is to consider the function defined for $r > 0$ by

$$h_r(x^m, \varphi; \theta) := \mathbf{1}_{[0,r]}(d(x^m, \varphi))\, e^{V(x^m|\varphi;\theta)}$$

where $d(x^m, \varphi) = \inf_{y^m \in \varphi} \|y - x\|$. Considering this function leads to

$$R_\Lambda(\varphi; h_r, \widehat{\theta}) = \int_{\Lambda \times \mathbb{M}} \mathbf{1}_{[0,r]}(d(x^m, \varphi))\mu(dx^m) - \sum_{x^m \in \varphi_\Lambda} h_r(x^m, \varphi \setminus x^m; \widehat{\theta}). \tag{5}$$

Then for a large window $R(\varphi; h_h, \widehat{\theta})/|\Lambda|$ leads to a difference of two estimates of the well-known empty space function F at distance r. Recall that for a marked stationary point process (see Møller and Waagepetersen (2003) for instance)

$$F(r) := P\left(d(0^M, \Phi) \le r\right).$$

The first term in the right hand side of (5) corresponds to the natural nonparametric estimator of $F(r)$ while the second one is a parametric estimator of $F(r)$.

3 Asymptotic properties

From now on, we assume that the point process satisfies **[Mod-E]** and **[Mod-L]**, that is **[Mod]**. The realization of $\Phi \sim P_{\theta^\star}$ is assumed to be observed in a domain $\Lambda_n \oplus D^+$, with $D^+ \ge D$, aimed at growing up to $\mathbb{R}^d$ as $n \to +\infty$. According to the locality assumption **[Mod-L]**, we are thus ensured that the $h-$innovations and $h-$residuals can be computed.

The aim of this section is to present several asymptotic properties for I_{Λ_n} and R_{Λ_n}. We prove their consistency and we propose two asymptotic normality results within different frameworks:

- Framework 1: for a fixed test function h, Λ_n is a cube, divided into a fixed finite number of sub-cubes (which will increase with Λ_n). The purpose is then to obtain the asymptotic normality for the vector composed of the $h-$residuals computed in each sub-cube.
- Framework 2: we consider $h_1, \dots, h_s$ s different test functions and the aim is to obtain the asymptotic normality of the vector composed of the h_j-residuals computed on Λ_n.

In both frameworks, an estimate of $\theta^\star$ is involved. We assume that it is computed from the full domain Λ_n with the same data used to evaluate the $h-$residuals, which is a natural setting in practice. Moreover, contrary to the previous works dealing with asymptotic properties on Gibbs point processes (*e.g.* Jensen and Künsch (1994), Comets and Janzura (1998) or Billiot et al. (2008)), where Λ_n is assumed to be of the discrete form $[-n, n]^d$, we consider general domains that may grow continuously up to $\mathbb{R}^d$.

The asymptotic results obtained in this section are the basis to derive goodness-of-fit tests, as presented in Section 5.

3.1 Consistency of the residuals process

We obtain consistency results for $I_{\tilde{\Lambda}_n}(\Phi; h, \theta^\star)$ and $R_{\tilde{\Lambda}_n}\left(\Phi; h, \widehat{\theta}_n(\Phi)\right)$, where for all $n \ge 1$, $\tilde{\Lambda}_n \subset \Lambda_n$, $(\tilde{\Lambda}_n)_{n\ge1}$ and $(\Lambda_n)_{n\ge1}$ are regular sequences whose size increases to ∞.

The assumption **[C]** gathers the two following assumptions:

[C1]

$$\mathbf{E}\left(\left|h\left(0^M,\Phi;\theta^\star\right)\right|e^{-V\left(0^M|\Phi;\theta^\star\right)}\right)<+\infty.$$

[C2] For all $(m,\varphi)\in\mathbb{M}\times\Omega$, the functions $h\left(0^m,\varphi;\theta\right)$ and $f\left(0^m,\varphi;\theta\right):=h\left(0^m,\varphi;\theta\right)e^{-V\left(0^m|\varphi;\theta\right)}$ are continuously differentiable with respect to θ in a neighborhood $\mathcal{V}(\theta^\star)$ of $\theta^\star$ and

$$\mathbf{E}\left(\left\|\mathbf{f}^{(1)}\left(0^M,\Phi;\theta^\star\right)\right\|\right)<+\infty\quad\text{ and }\quad\mathbf{E}\left(\left\|\mathbf{h}^{(1)}\left(0^M,\Phi;\theta^\star\right)\right\|e^{-V\left(0^M|\Phi;\theta^\star\right)}\right)<+\infty,$$

where $\mathbf{f}^{(1)}$ denotes the gradient vector of f with respect to θ.

Concerning the residuals process, we also need to assume

[E1] The estimator $\widehat{\theta}_n(\varphi)$ of $\theta^\star$, computed from the full observation domain Λ_n, converges for $P_{\theta^\star}-$a.e. φ towards $\theta^\star$, as $n\to+\infty$.

Proposition 2. *Assuming **[Mod]**, we have as $n\to+\infty$*

*(a) Under **[C1]**: for $P_{\theta^\star}-$a.e. φ, $|\tilde{\Lambda}_n|^{-1}I_{\tilde{\Lambda}_n}\left(\varphi;h,\theta^\star\right)$ converges towards 0.*

*(b) Under **[C]** and **[E1]**: for $P_{\theta^\star}-$a.e. φ, $|\tilde{\Lambda}_n|^{-1}R_{\tilde{\Lambda}_n}\left(\varphi;h,\widehat{\theta}_n(\varphi)\right)$ converges towards 0.*

Remark 1. *Assumption **[Mod-L]**, while useful to allow the computation of the residuals in practice, is actually useless to prove their consistency.*

3.2 Asymptotic control in probability of the residuals process

We provide in this section a control for the departure of the residuals from the innovations and $(\widehat{\theta}_n-\theta^\star)$. This is a crucial result to investigate the asymptotic normality of the residuals. We need the folllowing assumptions.

[N1] For all $(m,\varphi)\in\mathbb{M}\times\Omega$, the functions $h\left(0^m,\varphi;\theta\right)$ and $f\left(0^m,\varphi;\theta\right)$ (defined in **[C1]**) are twice continuously differentiable with respect to θ in a neighborhood $\mathcal{V}(\theta^\star)$ of $\theta^\star$ and

$$\mathbf{E}\left(\left\|\underline{\mathbf{f}}^{(2)}\left(0^M,\Phi;\theta^\star\right)\right\|\right)<+\infty\quad\text{ and }\quad\mathbf{E}\left(\left\|\underline{\mathbf{h}}^{(2)}\left(0^M,\Phi;\theta^\star\right)\right\|e^{-V\left(0^M|\Phi;\theta^\star\right)}\right)<+\infty,$$

where $\underline{\mathbf{g}}^{(2)}\left(0^m,\varphi;\theta^\star\right)=\left(\frac{\partial^2}{\partial\theta_j\partial\theta_k}g\left(0^m,\varphi;\theta^\star\right)\right)_{1\leq j,k\leq p}$ for $g=f,h$.

[E2] There exists a random vector $\mathbf{T}$ such that the following convergence holds as $n\to+\infty$

$$|\Lambda_n|^{1/2}\left(\widehat{\theta}_n(\Phi)-\theta^\star\right)\xrightarrow{d}\mathbf{T}.$$

Proposition 3. *Under assumptions **[C]**, **[N1]** and **[E1-2]**, assuming that $|\tilde{\Lambda}_n|=\mathcal{O}(|\Lambda_n|)$, then as $n\to+\infty$,*

$$R_{\tilde{\Lambda}_n}\left(\Phi;h,\widehat{\theta}_n(\Phi)\right)=I_{\tilde{\Lambda}_n}\left(\Phi;h,\theta^\star\right)-|\tilde{\Lambda}_n|\left(\widehat{\theta}_n(\Phi)-\theta^\star\right)^T\mathcal{E}\left(h;\theta^\star\right)+o_P(|\tilde{\Lambda}_n|^{1/2}),\qquad(6)$$

where $\mathcal{E}\left(h;\theta^\star\right)$ is the vector defined by

$$\mathcal{E}\left(h;\theta^\star\right):=\mathbf{E}\left(h\left(0^M,\Phi;\theta^\star\right)\mathbf{V}^{(1)}\left(0^M|\Phi;\theta^\star\right)e^{-V\left(0^M|\Phi;\theta^\star\right)}\right).\qquad(7)$$

The notation $X_n(\Phi)=o_P(w_n)$ means that $w_n^{-1}X_n(\Phi)$ converges in probability towards 0 as n tends to infinity.

Remark 2. *Note that for exponential family models, $\mathbf{V}^{(1)}(x^m|\varphi;\theta^\star)$ corresponds to the vector of sufficient statistics (see Section 6 for more details).*

3.3 Assumptions required for the asymptotic normality results

Apart from the assumptions **[Mod]**, **[C]** and **[N1]** on the model, we will need to assume **[N2-4]** below. All these assumptions are fulfilled by many models as proved in Section 6.

[N2] For any bounded domain Λ, for any $\theta \in \mathcal{V}(\theta^\star)$,

$$\mathbf{E}\left(\left| I_\Lambda\left(\Phi; h, \theta^\star\right)\right|^3 \right) < +\infty.$$

[N3] For any sequence of bounded domains Γ_n such that $\Gamma_n \to 0$ when $n \to \infty$, for any $\theta \in \mathcal{V}(\theta^\star)$,

$$\mathbf{E}\left(I_{\Gamma_n}\left(\Phi; h, \theta\right)^2 \right) \longrightarrow 0.$$

[N4] For any $\varphi \in \Omega$ and any bounded domain Λ, $I_\Lambda\left(\varphi; \theta\right)$ depends only on $\varphi_{\Lambda \oplus D}$.

Concerning the properties required for the estimator $\widehat{\theta}_n$, we need its consistency through **[E1]** and to refine **[E2]** into **[E2(bis)]** below. Note that the maximum pseudolikelihood estimator satisfies these assumptions for many models (see section 6.2).

[E2(bis)] The estimate admits the following expansion

$$\widehat{\theta}_n(\Phi) - \theta^\star = \frac{1}{|\Lambda_n|} \mathbf{U}_{\Lambda_n}\left(\Phi; \theta^\star\right) + o_P(|\Lambda_n|^{-1/2}),$$

where, for any $\theta \in \mathcal{V}(\theta^\star)$,

(i) for any $\varphi \in \Omega$ and for two disjoint bounded domains Λ_1, Λ_2,

$$\mathbf{U}_{\Lambda_1 \cup \Lambda_2}\left(\varphi; \theta\right) = \mathbf{U}_{\Lambda_1}\left(\varphi; \theta\right) + \mathbf{U}_{\Lambda_2}\left(\varphi; \theta\right),$$

(ii) for all $j = 1, \ldots, p$ and any bounded domain Λ

$$\mathbf{E}\left(\left| \left(\mathbf{U}_\Lambda\left(\Phi; \theta\right)\right)_j \right|^3 \right) < +\infty,$$

(iii) for all $j = 1, \ldots, p$ and for any bounded domain Λ

$$\mathbf{E}\left(\left(\mathbf{U}_\Lambda\left(\Phi; \theta\right)\right)_j \middle| \Phi_{\Lambda^c} \right) = 0,$$

(iv) for all $j = 1, \ldots, p$ and for any sequence of bounded domains Γ_n,

$$\mathbf{E}\left(\left(\mathbf{U}_{\Gamma_n}\left(\Phi; \theta\right)\right)_j^2 \right) \longrightarrow 0 \quad \text{as} \quad \Gamma_n \to 0,$$

(v) for any $\varphi \in \Omega$ and any bounded domain Λ, $\mathbf{U}_\Lambda\left(\varphi; \theta\right)$ depends only on $\varphi_{\Lambda \oplus D}$.

Remark 3. *Assumption **[E2(bis)]** implies **[E2]**. Indeed, under this assumption one may apply Theorem 2.1 of Jensen and Künsch (1994) and assert: there exists a matrix $\underline{\mathbf{\Sigma}}$ such that $|\Lambda_n|^{-1/2} \mathbf{U}_{\Lambda_n}\left(\Phi; \theta^\star\right) \xrightarrow{d} \mathcal{N}(0, \underline{\mathbf{\Sigma}})$, as $n \to +\infty$.*

3.4 Asymptotic normality of the $h-$residuals computed on subdomains of Λ_n

In this framework, we give ourself a test function h and we compute the $h-$residuals on disjoint subdomains of Λ_n. In this context, we assume that the domain Λ_n is a cube and is divided into a fixed number of subdomains as follows

$$\Lambda_n := \bigcup_{j\in\mathcal{J}} \Lambda_{j,n}$$

where $\mathcal{J}$ is a finite set and all the $\Lambda_{j,n}$ are disjoint cubes with the same volume $|\Lambda_{0,n}|$ increasing up to $+\infty$. Let us denote by $\mathbf{R}_{\mathcal{J},n}\left(\varphi; h, \widehat{\theta}_n\right)$ the vector of the residuals computed on each subdomain, i.e. $\mathbf{R}_{\mathcal{J},n}\left(\varphi; h, \widehat{\theta}_n\right) = \left(R_{\Lambda_{j,n}}\left(\varphi; h, \widehat{\theta}_n\right)\right)_{j\in\mathcal{J}}$.

According to Proposition 3 and in view of **[E2(bis)]**, we introduce the following notation

$$R_{\infty,\Lambda}(\varphi; h, \theta) := I_\Lambda(\varphi; h, \theta) - \mathbf{U}_\Lambda(\varphi;\theta)^T \mathcal{E}(h;\theta) \tag{8}$$

for any $\varphi \in \Omega$, for any bounded domain Λ and for any $\theta \in \boldsymbol{\Theta}$.

Proposition 4. *Assume that*

- *The parametric model satisfies* ***[Mod]****.*
- *The energy function and the test function h satisfy* ***[C]*** *and* ***[N1-4]****.*
- *The energy function and the estimate $\widehat{\theta}_n$ satisfy* ***[E1]*** *and* ***[E2(bis)]****.*

Then, the following convergence in distribution holds, as $n \to +\infty$

$$|\Lambda_{0,n}|^{-1/2}\, \mathbf{R}_{\mathcal{J},n}\left(\Phi; h, \widehat{\theta}_n\right) \xrightarrow{d} \mathcal{N}\left(0, \underline{\boldsymbol{\Sigma}}_1(\theta^\star)\right), \tag{9}$$

where $\underline{\boldsymbol{\Sigma}}_1(\theta^\star) = \lambda_{Inn}\,\underline{\mathbf{I}}_{|\mathcal{J}|} + |\mathcal{J}|^{-1}(\lambda_{Res} - \lambda_{Inn})\,\underline{\mathbf{J}}$ with $\underline{\mathbf{J}} = \mathbf{e}\mathbf{e}^T$ and $\mathbf{e} = (1, \ldots, 1)^T$. The constants λ_{Inn} and λ_{Res} are respectively defined by

$$\lambda_{Inn} = D^{-d} \sum_{|k|\le 1} \mathbf{E}\left(I_{\Delta_0(D)}(\Phi; h, \theta^\star)\, I_{\Delta_k(D)}(\Phi; h, \theta^\star)\right), \tag{10}$$

$$\lambda_{Res} = D^{-d} \sum_{|k|\le 1} \mathbf{E}\left(R_{\infty,\Delta_0(D)}(\Phi; h, \theta^\star) R_{\infty,\Delta_k(D)}(\Phi; h, \theta^\star)\right), \tag{11}$$

where, for all $k \in \mathbb{Z}^d$, $\Delta_k(D)$ is the cube centered at kD with side-length D.

From this asymptotic normality result, we can deduce the convergence for the norm of the centered residuals. This is the basis for a generalization of the quadrat counting test discussed in Section 5. We denote by $\overline{\mathbf{R}}_{\mathcal{J},n}(\varphi; h)$ the mean residuals over all subdomains, that is $\overline{\mathbf{R}}_{\mathcal{J},n}(\varphi; h) = |\mathcal{J}|^{-1} \sum_{j\in\mathcal{J}} R_{\Lambda_{j,n}}\left(\varphi; h, \widehat{\theta}_n\right)$.

Corollary 5. *Under the assumptions of Proposition 4,*

$$|\Lambda_{0,n}|^{-1} \|\mathbf{R}_{\mathcal{J},n}(\Phi; h) - \overline{\mathbf{R}}_{\mathcal{J},n}(\Phi; h)\|^2 \xrightarrow{d} \lambda_{Inn}\ \chi^2_{|\mathcal{J}|-1}. \tag{12}$$

Proof. An easy computation shows that λ_{Inn} and λ_{Res} are the two eigenvalues of $\underline{\boldsymbol{\Sigma}}_1(\theta^\star)$ with respective order $|\mathcal{J}| - 1$ and 1. Let $\underline{\mathbf{P}}_{Inn}$ be the matrix of orthonormalized eigenvectors associated to λ_{Inn}. This matrix of size $(|\mathcal{J}|, |\mathcal{J}| - 1)$ satisfies by definition $\underline{\mathbf{P}}_{Inn}^T \underline{\mathbf{P}}_{Inn} = \underline{\mathbf{I}}_{|\mathcal{J}|-1}$ and, from (9), $|\Lambda_{0,n}|^{-1}\ \|\underline{\mathbf{P}}_{Inn}^T \mathbf{R}_{\mathcal{J},n}(\varphi; h)\|^2 \xrightarrow{d} \lambda_{Inn}\ \chi^2_{|\mathcal{J}|-1}$. Moreover, it is easy to check that $\underline{\mathbf{P}}_{Inn}\underline{\mathbf{P}}_{Inn}^T = \underline{\mathbf{I}}_{|\mathcal{J}|} - |\mathcal{J}|^{-1}\underline{\mathbf{J}}_{|\mathcal{J}|}$ which leads to $\|\underline{\mathbf{P}}_{Inn}^T \mathbf{R}_{\mathcal{J},n}(\varphi; h)\|^2 = \|\mathbf{R}_{\mathcal{J},n}(\varphi; h) - \overline{\mathbf{R}}_{\mathcal{J},n}(\varphi; h)\|^2$. ■

Remark 4. *The asymptotic covariance matrix* $\underline{\mathbf{\Sigma}}_1(\theta^\star)$ *and* λ_{Inn} *involve only the covariance structure of the innovations (or the residuals) in a finite box around* 0*. This comes from the locality assumption [**Mod-L**], also involved in [**N4**] and [**E2(bis)**]. A challenging task in practice is to estimate* λ_{Inn} *and* λ_{Res} *(and so* $\underline{\mathbf{\Sigma}}_1(\theta^\star)$*), this issue is investigated in Section 4.*

3.5 Asymptotic normality of the $(h_j)_{j=1,\ldots,s}$−residuals computed on Λ_n

In this framework, we consider s different test functions and we compute all h_j−residuals on the same domain Λ_n, which is assumed to be a cube growing up to $\mathbb{R}^d$ when $n \to +\infty$.

We present an asymptotic normality result for the random vector $\left(R_{\Lambda_n}\left(\Phi; h_j, \widehat{\theta}_n\right)\right)_{j=1,\ldots,s}$.

Proposition 6. *Assume that*

- *The parametric model satisfies [**Mod**].*
- *The energy function and the test functions* h_j *(for* $j = 1, \ldots, s$*) satisfy [**C**] and [**N1-4**].*
- *The energy function and the estimate* $\widehat{\theta}_n$ *satisfy [**E1**] and [**E2(bis)**].*

Then, the following convergence in distribution holds, as $n \to +\infty$

$$|\Lambda_n|^{-1/2}\left(R_{\Lambda_n}\left(\Phi; h_j, \widehat{\theta}_n\right)\right)_{j=1,\ldots,s} \xrightarrow{d} \mathcal{N}\left(0, \underline{\mathbf{\Sigma}}_2(\theta^\star)\right), \tag{13}$$

where $\underline{\mathbf{\Sigma}}_2(\theta^\star)$ *is the* (s,s) *matrix given by*

$$\underline{\mathbf{\Sigma}}_2(\theta^\star) = D^{-d}\sum_{|k|\le 1}\mathbf{E}\left(\mathbf{R}_{\infty,\Delta_0(D)}(\Phi;\mathbf{h},\theta^\star)\,\mathbf{R}_{\infty,\Delta_k(D)}(\Phi;\mathbf{h},\theta^\star)^T\right), \tag{14}$$

where $\mathbf{R}_{\infty,\Lambda}(\varphi,\mathbf{h},\theta^\star) := (R_{\infty,\Lambda}(\varphi; h_j,\theta^\star))_{j=1,\ldots,s}$*, see (8), and where, for all* $k \in \mathbb{Z}^d$*,* $\Delta_k(D)$ *is the cube centered at* kD *with side-length* D*.*

4 Estimation and positivity of the asymptotic covariance matrices

4.1 Statement of the problem

The aim of this section is to provide a condition under which, on the one hand the matrices $\underline{\mathbf{\Sigma}}_1(\theta^\star)$ and $\underline{\mathbf{\Sigma}}_2(\theta^\star)$, defined in Propositions 4 and 6, are positive-definite, and on the other hand λ_{Inn}, involved in Corollary 5, is positive. Then we define estimators of $\underline{\mathbf{\Sigma}}_1^{-1/2}(\theta^\star)$, λ_{Inn}^{-1} and $\underline{\mathbf{\Sigma}}_2^{-1/2}(\theta^\star)$. As a consequence, we will be in position to normalize and estimate the quantities arising in (9), (12) and (13) so that they converge to a free law.

Before this, let us focus on the particular form of the matrix $\underline{\mathbf{\Sigma}}_1(\theta^\star)$. This $(|\mathcal{J}|,|\mathcal{J}|)$ matrix has two eigenvalues λ_{Inn} and λ_{Res} (respectively defined by (10) and (11)), whose multiplicity is $|\mathcal{J}|-1$ for λ_{Inn} and 1 for λ_{Res}. By using the Gram-Schmidt process for orthonormalizing the eigenvectors of $\underline{\mathbf{\Sigma}}_1(\theta^\star)$, one obtains the explicit form for the squared inverse of this matrix, provided λ_{Inn} and λ_{Res} do not vanish:

$$\underline{\mathbf{\Sigma}}_1^{-1/2}(\theta^\star) = \frac{1}{\sqrt{\lambda_{Inn}}}\,\underline{\mathbf{I}}_{|\mathcal{J}|} + \frac{1}{|\mathcal{J}|}\left(\frac{1}{\sqrt{\lambda_{Res}}} - \frac{1}{\sqrt{\lambda_{Inn}}}\right)\underline{\mathbf{J}},$$

where $\underline{\mathbf{J}} = \mathbf{e}\mathbf{e}^T$ and $\mathbf{e} = (1,\ldots,1)^T$. Therefore, estimating $\underline{\mathbf{\Sigma}}_1^{-1/2}(\theta^\star)$ can be reduced to the estimation of these two eigenvalues λ_{Inn} and λ_{Res}.

Consequently, the estimation of λ_{Inn} and the covariance matrices $\underline{\mathbf{\Sigma}}_1(\theta^\star)$ and $\underline{\mathbf{\Sigma}}_2(\theta^\star)$ is achieved by estimating (10), (11) and (14), which can be viewed as a particular case of estimating the matrix (actually a constant for the two first expressions)

$$\underline{\mathbf{M}}(\theta^\star) = D^{-d} \sum_{|k|\leq 1} \mathbf{E}\left(\mathbf{Y}_{\Delta_0(D)}(\Phi;\theta^\star)\, \mathbf{Y}_{\Delta_k(D)}(\Phi;\theta^\star)^T \right),$$

where, according to the assumptions involved in Propositions 4 and 6, for any bounded domain Λ, $\mathbf{Y}_\Lambda(\Phi;\theta)$ is a random vector of dimension q ($q=1$ or s) depending on θ, such that for any bounded domains $\Lambda, \Lambda_1, \Lambda_2$ (Λ_1, Λ_2 disjoint), for any $\theta \in \mathcal{V}(\theta^\star)$, for any $j=1,\ldots,q$ and any $\varphi \in \Omega$

(i) $\mathbf{Y}_{\Lambda_1\cup\Lambda_2}(\varphi;\theta) = \mathbf{Y}_{\Lambda_1}(\varphi;\theta) + \mathbf{Y}_{\Lambda_2}(\varphi;\theta)$,

(ii) $\mathbf{E}\left((\mathbf{Y}_\Lambda(\Phi;\theta))_j^2 \right) < +\infty$,

(iii) $\mathbf{E}\left((\mathbf{Y}_\Lambda(\Phi;\theta))_j \,\middle|\, \Phi_{\Lambda^c} \right) = 0$,

(iv) for any sequence of bounded domains Γ_n, $\mathbf{E}\left((\mathbf{Y}_{\Gamma_n}(\Phi;\theta))_j^2 \right) \longrightarrow 0 \quad \text{as} \quad \Gamma_n \to 0$,

(v) $\mathbf{Y}_\Lambda(\varphi;\theta)$ depends only on $\varphi_{\Lambda\oplus D}$.

4.2 Positive definiteness of $\underline{\mathbf{M}}(\theta^\star)$

Let us consider the following assumption.

[PD] For some $\overline{\Lambda} := \cup_{|i|\leq\lceil \frac{D}{\overline{\delta}}\rceil} \Delta_i(\overline{\delta})$ with $\overline{\delta} > 0$, there exists $B \in \mathcal{F}$ and $A_0, \ldots, A_\ell$, ($\ell \geq 1$) disjoint events of $\overline{\Omega}_B := \left\{ \varphi \in \Omega : \varphi_{\Delta_i(\overline{\delta})} \in B, 1 \leq |i| \leq 2\left\lceil \frac{D}{\overline{\delta}} \right\rceil \right\}$ such that

- for $j = 0, \ldots, \ell$, $P_{\theta^\star}(A_j) > 0$.
- for all $(\varphi_0, \ldots, \varphi_\ell) \in A_0 \times \cdots \times A_\ell$ the (ℓ, q) matrix with entries $(\mathbf{Y}_{\overline{\Lambda}}(\varphi_i;\theta^\star))_j - (\mathbf{Y}_{\overline{\Lambda}}(\varphi_0;\theta^\star))_j$ is injective, which means:

$$\left(\forall \mathbf{y} \in \mathbb{R}^q, \mathbf{y}^T \left(\mathbf{Y}_{\overline{\Lambda}}(\varphi_i;\theta^\star) - \mathbf{Y}_{\overline{\Lambda}}(\varphi_0;\theta^\star) \right) = 0 \right) \Longrightarrow \mathbf{y} = 0.$$

Proposition 7. *From the definition of $\mathbf{Y}_\Lambda(\Phi;\theta)$ and under **[PD]**, the matrix $\underline{\mathbf{M}}(\theta^\star)$ is positive-definite.*

Remark 5. *The assumption **[PD]** is associated to some characteristics of the point process Φ. The parameter $\overline{\delta}$ is independent of the parameters involved in the different estimators (e.g. $D^\vee$ or δ arising in the next section). Given a model, the event B and $\overline{\delta}$ are chosen in order to let the different configurations sets $A_0, A_1, \ldots, A_\ell$ as simple as possible. For most models, a convenient choice is $B = \emptyset$ and $\overline{\delta} \geq D$ (see the examples treated in Appendix B for instance).*

4.3 Estimation of $\underline{\mathbf{M}}(\theta^\star)$

The dependence of $\underline{\mathbf{M}}(\theta^\star)$ on D may be lightened thanks to the following lemma, whose proof is relegated to section 8.5.

Lemma 8. *The matrix $\underline{\mathbf{M}}(\theta^\star)$ can be rewritten for any $\delta > 0$ and any $D^\vee \geq D$ as*

$$\underline{\mathbf{M}}(\theta^\star) = \delta^{-d} \sum_{|k|\leq\lceil \frac{D^\vee}{\delta}\rceil} \mathbf{E}\left(\mathbf{Y}_{\Delta_0(\delta)}(\Phi;\theta^\star)\, \mathbf{Y}_{\Delta_k(\delta)}(\Phi;\theta^\star)^T \right),$$

where $\Delta_k(\delta)$ is the cube with side-length δ centered at $k\delta$.

From this result, to achieve an estimation of $\underline{\mathbf{M}}(\theta^\star)$, it is required to estimate the involved expectation and $\theta^\star$ (by $\widehat{\theta}_n$). This is enough for the estimation of λ_{Inn} for which $\mathbf{Y}_\Lambda(\varphi;\theta) = I_\Lambda(\varphi;\theta)$. But when $\mathbf{Y}_\Lambda(\varphi;\theta) = R_{\infty,\Lambda}(\varphi;h,\theta)$ or $\mathbf{Y}_\Lambda(\varphi;\theta) = \mathbf{R}_{\infty,\Lambda}(\varphi;\mathbf{h},\theta)$, which appears in $\underline{\mathbf{\Sigma}}_1(\theta^\star)$ and $\underline{\mathbf{\Sigma}}_2(\theta^\star)$, it can be noticed that $\mathbf{Y}_\Lambda$ still depends on an expectation with respect to $P_{\theta^\star}$, through the vector $\mathcal{E}(h,\theta^\star)$ defined by (7). Moreover, the vector $\mathbf{U}_\Lambda$ in **[E2(bis)]** may also depend on such a term (this is the case for example when considering the maximum pseudolikelihood estimate as shown in Section 6.2). This means that $\mathbf{Y}_\Lambda(\varphi;\theta)$ cannot be estimated only by $\mathbf{Y}_\Lambda(\varphi;\widehat{\theta}_n)$, but by $\widehat{\mathbf{Y}}_{n,\Lambda}(\varphi;\widehat{\theta}_n)$, where $\widehat{\mathbf{Y}}_n$ is an estimator of $\mathbf{Y}$. We assume in the sequel that $\widehat{\mathbf{Y}}_n$ satisfies the same properties $(i)-(v)$ as $\mathbf{Y}$ and is a good estimator of $\mathbf{Y}$ (see Proposition 9). The explicit form of $\widehat{\mathbf{Y}}_n$ depends strongly on the estimate $\widehat{\theta}_n$ (e.g. through $\mathbf{U}_\Lambda$ in **[E2(bis)]**). When $\widehat{\theta}_n$ is the maximum pseudolikelihood estimator, we provide explicit formulas for $\widehat{\mathbf{Y}}_n$ in Section 6.3.

Let us now specify an estimator of $\underline{\mathbf{M}}(\theta^\star)$. Assume that the point process is observed in the domain $\Lambda_{n_0} \oplus D^+$ where $D^+ \geq D$ and Λ_{n_0} is a cube. For any δ such that $|\Lambda_{n_0}|\delta^{-d} \in \mathbb{N}$, we may consider the decomposition $\Lambda_{n_0} = \cup_{k\in\mathcal{K}_{n_0}} \Delta_k(\delta)$, where the $\Delta_k(\delta)$'s are disjoint cubes with side-length δ and centered, without loss of generality, at $k\delta$. For any such δ, according to Lemma 8, a natural estimator of $\underline{\mathbf{M}}(\theta^\star)$ is, for any $D^\vee \geq D$,

$$\widehat{\underline{\mathbf{M}}}_{n_0}(\varphi;\widehat{\theta}_{n_0}(\varphi),\delta,D^\vee) = |\Lambda_{n_0}|^{-1} \sum_{k\in\mathcal{K}_{n_0}} \sum_{j\in\mathbb{B}_k\left(\lceil \frac{D^\vee}{\delta}\rceil\right)\cap\mathcal{K}_{n_0}} \widehat{\mathbf{Y}}_{n_0,\Delta_j(\delta)}\left(\varphi;\widehat{\theta}_{n_0}(\varphi)\right) \widehat{\mathbf{Y}}_{n_0,\Delta_k(\delta)}\left(\varphi;\widehat{\theta}_{n_0}(\varphi)\right)^T. \tag{15}$$

Remark 6. *As suggested by Lemma 8, the parameter δ in (15) may be chosen arbitrarily. Yet, while $\underline{\mathbf{M}}(\theta^\star)$ is actually independent of δ, its estimate $\widehat{\underline{\mathbf{M}}}_{n_0}$ may depend on it due to edge effects.*

The following proposition provides a framework to study the asymptotic properties of (15) and shows the consistency of $\widehat{\underline{\mathbf{M}}}_{n_0}$ when the domain Λ_n increases up to ∞ as $n\to\infty$. Its proof is relegated to section 8.7.

Proposition 9. *Under **[Mod]**, **[E1]**, assume that for any θ in a neighborhood $\mathcal{V}(\theta^\star)$ of $\theta^\star$, for any bounded domain Λ, for any $\varphi\in\Omega$ and for $j=1,\ldots,p$, $\left(\widehat{\mathbf{Y}}_{n,\Lambda}(\varphi;\cdot)\right)_j$ is a continuous function. Assume moreover that*

$$\sup_{k\in\mathcal{K}_n} \left|\widehat{\mathbf{Y}}_{n,\Delta_k(\delta_n)}(\Phi;\theta) - \mathbf{Y}_{\Delta_k(\delta_n)}(\Phi;\theta)\right| \xrightarrow{\mathbb{P}} 0, \tag{16}$$

where, for any $\delta>0$ as above, $(\delta_n)_{n\in\mathbb{N}}$ is a sequence satisfying $|\Lambda_n|\delta_n^{-d}\in\mathbb{N}$, $\delta_{n_0}=\delta$ and $\delta_n\to\delta$ as $n\to\infty$. Then, for any $D^\vee\geq D$,

$$\widehat{\underline{\mathbf{M}}}_n\left(\Phi;\widehat{\theta}_n(\Phi),\delta_n,D^\vee\right) \xrightarrow{\mathbb{P}} \underline{\mathbf{M}}(\theta^\star).$$

Remark 7. *The choice of the sequence $(\delta_n)_{n\in\mathbb{N}}$ is always possible (see the proof). Since we allow the domain Λ_n to grow continuously up to $\mathbb{R}^d$, its decomposition in sub-cubes with side-length δ is not always possible. The sequence $(\delta_n)_{n\in\mathbb{N}}$ is thus mandatory to make a decomposition of the domain available when n increases. We chose it by respecting as most as possible the initial choice of the practicioner.*

5 Goodness-of-fit tests for stationary marked Gibbs point processes

We present in this section three goodness-of-fit tests, based on the residuals computed according to the different frameworks considered in Section 3. We assume that the point process is observed in the domain $\Lambda_{n_0}\oplus D^+$ where $D^+\geq D$ and Λ_{n_0} is a cube.

5.1 Quadrat-type test with $|J|-1$ degrees of freedom

According to the setting detailed in Section 3.4, we divide the domain Λ_{n_0} into a fixed number of subdomains, namely $\Lambda_{n_0} := \bigcup_{j\in\mathcal{J}} \Lambda_{j,n_0}$ where $\mathcal{J}$ is a finite set and all the Λ_{j,n_0} are disjoint cubes with the same volume $|\Lambda_{0,n_0}|$. Moreover, in each sub-domain, we consider the decomposition $\Lambda_{j,n_0} = \cup_{k\in\mathcal{K}_{j,n_0}} \Delta_k(\delta)$, for any δ such that $|\Lambda_{0,n_0}|\delta^{-d} \in \mathbb{N}$, where the $\Delta_k(\delta)$'s are disjoint cubes with side-length δ.

Following (15), we consider, for any $\delta > 0$ as above and any $D^\vee \geq D$, the estimator

$$\widehat{\lambda}_{n_0,Inn} = |\Lambda_{n_0}|^{-1} \sum_{i\in\mathcal{K}_{n_0}} \sum_{j\in\mathbb{B}_i(\lceil \frac{D^\vee}{\delta}\rceil)\cap\mathcal{K}_{n_0}} I_{\Delta_i(\delta)}\left(\varphi;\widehat{\theta}_{n_0}(\varphi)\right) I_{\Delta_j(\delta)}\left(\varphi;\widehat{\theta}_{n_0}(\varphi)\right), \tag{17}$$

where $\mathcal{K}_{n_0} = \cup_{j\in\mathcal{J}}\mathcal{K}_{j,n_0}$. Note that $I_{\Delta_i(\delta)}\left(\varphi;\widehat{\theta}_{n_0}(\varphi)\right) = R_{\Delta_i(\delta)}\left(\varphi;\widehat{\theta}_{n_0}(\varphi)\right)$ but we preserve this redundant notation in the sequel.

The following corollary is an immediate consequence of Corollary 5 and Proposition 9.

Corollary 10. *Under the assumptions of Proposition 4 and if **[PD]** holds for $\mathbf{Y}_{\overline{\Lambda}}(\Phi;\theta^\star) = I_{\overline{\Lambda}}(\Phi;\theta^\star)$, then, for any $\delta > 0$, one can construct a sequence $(\delta_n)_{n\in\mathbb{N}}$ satisfying $|\Lambda_{0,n}|\delta_n^{-d} \in \mathbb{N}$, $\delta_{n_0} = \delta$ and $\delta_n \to \delta$, such that as $n \to +\infty$*

$$T_{1,n} := |\Lambda_{0,n}|^{-1}\, \widehat{\lambda}_{n,Inn}^{-1} \times \|\mathbf{R}_{\mathcal{J},n}(\Phi;h) - \overline{\mathbf{R}}_{\mathcal{J},n}(\Phi;h)\|^2 \xrightarrow{d} \chi^2(|\mathcal{J}|-1). \tag{18}$$

This result leads to a goodness-of-fit test for $H_0 : \Phi \sim P_{\theta^\star}$ versus $H_1 : \Phi \nsim P_{\theta^\star}$. Let us briefly summarize the different steps to implement the test for a given asymptotic level $\alpha \in (0,1)$.

- **Step 1** Consider a parametric model of a stationary marked Gibbs point process with finite range D observed on the domain $\Lambda_{n_0} \oplus D^+$ with $D^+ \geq D$.
- **Step 2** Choose an estimation method satisfying the assumptions **[E1]**, **[E2(bis)]** (for example the MPLE) and compute the estimate $\widehat{\theta}_{n_0}$ on Λ_{n_0}.
- **Step 3**
  a) Consider a test function h (satisfying **[C1-2]**, **[N1-3]** and **[PD]**), divide Λ_{n_0} into $|\mathcal{J}|$ cubes and compute the $h-$residuals on each different cube.
  b) Estimate λ_{Inn} by (17).
  c) Compute the test statistic T_{1,n_0} involved in (18).
- **Step 4** Reject the model if $T_{1,n_0}(\varphi) > \chi^2_{1-\alpha}(|\mathcal{J}|-1)$.

Let us note that in the particular case of a homogeneous Poisson point process with intensity z and when considering the raw residuals ($h=1$), this test is exactly the Poisson dispersion test applied to the $|\mathcal{J}|$ quadrat counts, also called quadrat counting test, see Diggle (2003) for instance. Indeed, in this case, $\mathbf{R}_{\mathcal{J},n}(\varphi;h) - \overline{\mathbf{R}}_{\mathcal{J},n}(\varphi;h)$ is the vector of quadrat counts and $\lambda_{Inn} = z$. Considering $|\Lambda_{0,n}|\widehat{\lambda}_{n_0,Inn}$ as an estimation of the intensity on $\Lambda_{0,n}$, the statistic $T_{1,n}$ reduces to the ratio of the sum of squares of the quadrat counts over their estimated mean.

Remark 8. *The condition **[PD]** in Corollary 10 has to be verified with $\mathbf{Y}_{\overline{\Lambda}}(\Phi;\theta^\star) = I_{\overline{\Lambda}}(\Phi;\theta^\star)$ which is not so difficult (see Proposition 16 for a general result). Indeed, contrarily to Corollary 11 and 12, this condition does not depend on the form of the estimator $\widehat{\theta}_n$. Moreover, as emphasized in Section 3.4, the assumptions of Proposition 4 are satisfied for many models (this will be explored in details for exponential models in Section 6.1). This means that the proposed goodness-of-fit test based on (18) may be used for many models and many choices of function h.*

Remark 9. *The weakeness of this testing procedure (and the next ones) could be the estimation (17) of λ_{Inn} (and in general the estimator (15)). The choice of the parameters δ and $D^\vee$ in (17) is crucial. For instance, for fixed n, in the extreme cases $\delta \to 0$ or $D^\vee \to \infty$, we get $\widehat{\lambda}_{n,Inn} \approx 0$. A careful simulation study should help for these choices. Another improvement could be to estimate λ_{Inn} via Monte-Carlo methods.*

5.2 Quadrat-type test with $|J|$ degrees of freedom

Under the same setting as above, assume moreover that **[PD]** holds for $\mathbf{Y}_{\overline{\Lambda}}(\varphi;\theta^\star) = R_{\infty,\overline{\Lambda}}(\varphi;h,\theta^\star)$. Let us define the normalized residuals

$$\widetilde{\mathbf{R}}_{1,n_0}(\varphi;h) := \widehat{\lambda}_{n_0,Inn}^{-1/2}\mathbf{R}_{\mathcal{J},n_0}(\varphi;h) + \left(\widehat{\lambda}_{n_0,Res}^{-1/2} - \widehat{\lambda}_{n_0,Inn}^{-1/2}\right)\overline{\mathbf{R}}_{\mathcal{J},n_0}(\varphi;h),$$

where $\widehat{\lambda}_{n_0,Inn}$ is defined in (17) and $\widehat{\lambda}_{n_0,Res}$ is an estimate of λ_{Res} following (15). When considering the MPLE, explicit formulas for $\widehat{\lambda}_{n_0,Res}$ are given in Section 6.3. It is easy to check that $\widetilde{\mathbf{R}}_{1,n_0}(\varphi;h) = \widehat{\underline{\mathbf{\Sigma}}_{1n_0}}^{-1/2} \mathbf{R}_{\mathcal{J},n_0}(\varphi;h)$. Therefore the following corollary is deduced from Propositions 4 and 9.

Corollary 11. *Under the assumptions of Propositions 4 and 9, assuming that **[PD]** holds for $\mathbf{Y}_{\overline{\Lambda}}(\Phi;\theta^\star) = I_{\overline{\Lambda}}(\Phi;\theta^\star)$ and $\mathbf{Y}_{\overline{\Lambda}}(\Phi;\theta^\star) = R_{\infty,\overline{\Lambda}}(\Phi;h,\theta^\star)$, then, for any $\delta > 0$, one can construct a sequence $(\delta_n)_{n\in\mathbb{N}}$ which satisfies $|\Lambda_{0,n}|\delta_n^{-d} \in \mathbb{N}$, $\delta_{n_0} = \delta$ and $\delta_n \to \delta$ as $n \to \infty$, such that as $n \to +\infty$,*

$$\widetilde{T}_{1,n}(\Phi) := |\Lambda_{0,n}|^{1/2}\|\widetilde{\mathbf{R}}_{1,n}(\Phi;h)\|^2 \xrightarrow{d} \chi^2(|\mathcal{J}|) \tag{19}$$

A goodness-of-fit test with asymptotic size $\alpha \in (0,1)$ is deduced similarly as in the previous section. The steps to follow in practice are the same except that in **Step 3** b), one has to estimate both λ_{Inn} and λ_{Res}, and in **Step 4** we reject the model if $\widetilde{T}_{1,n_0}(\varphi) > \chi^2_{1-\alpha}(|\mathcal{J}|)$.

Remark 10. *Let us emphasize that, with respect to Corollary 10, Corollary 11 involves an additional more complex assumption: **[PD]** has to be satisfied for $\mathbf{Y}_{\overline{\Lambda}}(\Phi;\theta^\star) = R_{\infty,\overline{\Lambda}}(\Phi;h,\theta^\star)$. This kind of assumption deeply depends on the nature of the estimate $\widehat{\theta}$. This problem is investigated in Proposition 18 for particular examples. Furthermore, we show in Proposition 17 that $\lambda_{Res} = 0$ occurs for many models and many choices of h including the Poisson model when $h = 1$. These two remarks underline the fact that the test relying on $\widetilde{T}_{1,n}$ is more restrictive than the previous one with $T_{1,n}$.*

5.3 Empty space function type test

Let us consider the setting of section 3.5, where s different residuals are computed on the same full domain Λ_{n_0}. We consider the decomposition $\Lambda_{n_0} = \cup_{k\in\mathcal{K}_{n_0}}\Delta_k(\delta)$, for any δ such that $|\Lambda_{n_0}|\delta^{-d} \in \mathbb{N}$, where the $\Delta_k(\delta)$'s are disjoint cubes with side-length δ.

Under the notation of Proposition 6, assuming **[PD]** holds for $\mathbf{Y}_{\overline{\Lambda}}(\varphi;\theta^\star) = \mathbf{R}_{\infty,\overline{\Lambda}}(\varphi,\mathbf{h},\theta^\star)$, let us define

$$\widetilde{\mathbf{R}}_{2,n_0}(\varphi;\mathbf{h},\widehat{\theta}) := \widehat{\underline{\mathbf{\Sigma}}_{2n_0}}^{-1/2} \left(R_{\Lambda_{n_0}}(\varphi;h_j,\widehat{\theta})\right)_{j=1,\dots,s}$$

where $\widehat{\underline{\mathbf{\Sigma}}_{2n_0}}^{-1/2} := \widehat{\underline{\mathbf{\Sigma}}_{2n_0}}^{-1/2}(\varphi,\widehat{\theta};\delta,D^\vee)$ is an estimation of $\underline{\mathbf{\Sigma}}_2(\theta^\star)$ as in (15). See explicit formulas in Section 6.3 when considering the MPLE.

From Propositions 6 and 9, we get the following corollary.

Corollary 12. *Assuming **[PD]** with $\mathbf{Y}_{\overline{\Lambda}}(\varphi;\theta^\star) = \mathbf{R}_{\infty,\overline{\Lambda}}(\varphi,\mathbf{h},\theta^\star)$, under the assumptions of Propositions 6 and 9, then, for any $\delta > 0$ as above, one can construct a sequence $(\delta_n)_{n\in\mathbb{N}}$ which satisfies $|\Lambda_n|\delta_n^{-d} \in \mathbb{N}$, $\delta_{n_0} = \delta$ and $\delta_n \to \delta$ as $n \to \infty$, such that, as $n \to +\infty$,*

$$\widetilde{T}_{2,n}(\Phi) := |\Lambda_n|^{1/2}\|\widetilde{\mathbf{R}}_{2,n}(\Phi;\mathbf{h},\widehat{\theta})\|^2 \xrightarrow{d} \chi^2(s). \tag{20}$$

A goodness-of-fit test for $H_0 : \Phi \sim P_{\theta^\star}$ versus $H_1 : \Phi \nsim P_{\theta^\star}$, with asymptotic size $\alpha \in (0,1)$ is deduced as before. From a practical point of view, the steps detailed in 5.1 are modified into:

- **Step 3(framework 2)**
  a) Consider s different test functions (satisfying **[C1-2]**, **[N1-3]** and **[PD]**), and compute the s different h_j-residuals on the same initial domain Λ_{n_0}.
  b) Estimate the matrix $\underline{\mathbf{\Sigma}}_2(\theta^\star)$ by (15) and compute $\widehat{\underline{\mathbf{\Sigma}}}_{2n_0}^{-1/2}$ with any numerical routine (e.g. a choleski decomposition or a singular value decomposition).
  c) Compute the test statistic $\widetilde{T}_{2,n_0}(\varphi)$ defined by (20).
- **Step 4** Fix $\alpha \in (0,1)$ and reject the model if $\widetilde{T}_{2,n_0}(\varphi) > \chi^2_{1-\alpha}(s)$.

6 Application to exponential models and the MPLE

Through Sections 3, 4 and 5 three sets of assumptions have been considered. The first one deals with integrability and regularity of the model and the test function(s) and gathers **[Mod]**, **[C]** and **[N1-4]**. The second one is about the estimator $\widehat{\theta}_n$ and involves **[E1]** and **[E2(bis)]**. Finally, the third one, assumption **[PD]** is very specific and deals with the positive definiteness of covariance matrices. We prove in this section that these assumptions are in general fulfilled for exponential family models and the MPLE.

6.1 Assumptions [Mod], [C] and [N1-4] for exponential family models

The energy function of exponential family models is given for any $\Lambda \in \mathcal{B}(\mathbb{R}^d)$ by $V_\Lambda(\varphi;\theta) = \theta^T \mathbf{v}_\Lambda(\varphi)$, where $\mathbf{v}_\Lambda(\varphi)$ is the vector of sufficient statistics given by $\mathbf{v}_\Lambda(\varphi) = (v_{1,\Lambda}(\varphi), \ldots, v_{p,\Lambda}(\varphi))^T$. The local energy is then expressed as $V(x^m|\varphi;\theta) = \theta^T \mathbf{v}(x^m|\varphi)$, where $\mathbf{v}(x^m|\varphi) = (v_1(x^m|\varphi), \ldots, v_p(x^m|\varphi)) := \mathbf{v}_\Lambda(\varphi \cup \{x^m\}) - \mathbf{v}_\Lambda(\varphi)$. Let us consider the following assumption:

[Exp] For $i = 1, \cdots, p$, there exist $\kappa_i^{(\text{inf})}, \kappa_i^{(\text{sup})} \geq 0$, $k_i \in \mathbb{N}$ such that one of both following assumptions is satisfied for all $(m,\varphi) \in \mathbb{M} \times \Omega$:

$$\theta_i \geq 0 \text{ and } -\kappa_i^{(\text{inf})} \leq v_i(0^m|\varphi) = v_i(0^m|\varphi_{\mathcal{B}(0,D)}) \leq \kappa_i^{(\text{sup})} |\varphi_{\mathcal{B}(0,D)}|^{k_i}.$$

or

$$-\kappa_i^{(\text{inf})} \leq v_i(0^m|\varphi) = v_i(0^m|\varphi_{\mathcal{B}(0,D)}) \leq \kappa_i^{(\text{sup})}.$$

The assumption **[Exp]** has already been considered by Billiot et al. (2008). It is fulfilled for a large class of examples including the overlap area point process, the multi-Strauss marked point process, the $k-$nearest-neighbor multi-Strauss marked point process, the Strauss type disc process, the Geyer's triplet point process, the area interaction point process,....

Proposition 13. *Under **[Exp]**, the assumptions **[Mod]**, **[C]** and **[N1-4]** are satisfied for the raw residuals, inverse residuals, Pearson residuals or residuals based on the empty space function.*

Proof. The assumption **[Exp]** implies that the local energy function is local and stable, which, from results of Bertin et al. (1999), implies that **[Mod]** is fulfilled. A direct consequence of **[Exp]** is that for every $\alpha > 0$, for all $\theta \in \mathbf{\Theta}$ and for all $i = 1, \ldots, p$

$$\mathbf{E}\left(|v_i(0^M|\Phi)|^\alpha e^{-\theta^T \mathbf{v}(0^M|\Phi)}\right) < +\infty, \tag{21}$$

which ensures the integrability assumptions **[C]** and **[N1-2]** for the residuals considered in the proposition. The locality assumption **[N4]** is contained in **[Exp]**. Finally, an application of the dominated convergence theorem, with the help of (21), shows **[N3]**. ■

Remark 11. *Our setting is not restricted to locally stable exponential family models. As an example, following ideas of Coeurjolly and Drouilhet (2009), one may prove that **[C]** and **[N1-4]** are fulfilled for Lennard-Jones type models.*

6.2 Assumptions [E1] and [E2(bis)] for the MPLE

Among the different parametric estimation methods available for spatial point patterns, the maximum pseudolikelihood is of particular interest. Indeed, unlike the maximum likelihood estimation method, it does not require the computation of the partition function, it is quite easy to implement and asymptotic results are now well-known (see Jensen and Møller (1991), Jensen and Künsch (1994), Billiot et al. (2008), Dereudre and Lavancier (2009) and Coeurjolly and Drouilhet (2009)). The MPLE is obtained by maximizing the log-pseudolikelihood contrast, given for exponential models by

$$LPL_{\Lambda_n}(\varphi;\theta) = -\int_{\Lambda_n\times\mathbb{M}} e^{-\theta^T\mathbf{v}(x^m|\varphi)}\mu(dx^m) \; - \; \theta^T\sum_{x\in\varphi}\mathbf{v}(x^m|\varphi\setminus x^m). \tag{22}$$

Proposition 14. *Under assumption **[Exp]** (and an additional indentifiability condition), **[E1]** and **[E2(bis)]** are fulfilled for the MPLE. The vector $\mathbf{U}_\Lambda(\varphi;\theta^\star)$ in **[E2(bis)]** is then expressed as follows*

$$\mathbf{U}_\Lambda(\varphi;\theta^\star) := \underline{\mathbf{H}}(\theta^\star)^{-1}\mathbf{LPL}_\Lambda^{(1)}(\varphi;\theta^\star), \tag{23}$$

where $\mathbf{LPL}_\Lambda^{(1)}(\varphi;\theta^\star)$ is the gradient vector of the log-pseudolikelihood given by

$$\mathbf{LPL}_\Lambda^{(1)}(\varphi;\theta^\star) := \int_{\Lambda\times\mathbb{M}}\mathbf{v}(x^m|\varphi;\theta^\star)e^{-\theta^{\star T}\mathbf{v}(x^m|\varphi)}\mu(dx^m) - \sum_{x^m\in\varphi_\Lambda}\mathbf{v}(x^m|\varphi\setminus x^m;\theta^\star) \tag{24}$$

and where $\underline{\mathbf{H}}(\theta^\star)$ is the symmetric matrix

$$\underline{\mathbf{H}}(\theta^\star) := \mathbf{E}\left(\mathbf{v}(0^M|\Phi;\theta^\star)\mathbf{v}(0^M|\Phi;\theta^\star)^T\; e^{-V\left(0^M|\Phi;\theta^\star\right)}\right). \tag{25}$$

Proof. **[E1]** is proved by Billiot et al. (2008) (under **[Exp]** and the identifiability condition **[Ident]** arising p.244 in Billiot et al. (2008)). Let $\mathbf{Z}_n(\varphi;\theta^\star) := -|\Lambda_n|^{-1}\mathbf{LPL}_{\Lambda_n}(\varphi;\theta^\star)$. If $\widehat{\theta}_n(\varphi) = \widehat{\theta}_n^{MPLE}(\varphi)$ denotes the maximum pseudolikelihood estimate, one derives

$$\mathbf{Z}_n^{(1)}(\varphi;\widehat{\theta}_n) - \mathbf{Z}_n^{(1)}(\varphi;\theta^\star) = 0 - \mathbf{Z}_n^{(1)}(\varphi;\theta^\star) = \underline{\mathbf{H}}_n(\varphi;\theta^\star,\widehat{\theta}_n)(\widehat{\theta}_n(\varphi)-\theta^\star)$$

with $\underline{\mathbf{H}}_n(\varphi;\theta^\star,\widehat{\theta}_n) = \int_0^t \underline{\mathbf{Z}}_n^{(2)}\left(\varphi;\theta^\star + t(\widehat{\theta}_n(\varphi)-\theta^\star)\right)dt$. Under assumptions **[Exp]** and **[Ident]**, then, for n large enough, $\underline{\mathbf{H}}_n$ is invertible and converges almost surely towards the matrix $\underline{\mathbf{H}}(\theta^\star)$ given by (25). Moreover, following the proof of Theorem 2 of Billiot et al. (2008) (see condition (iii) p.257-258), we derive $Var(\mathbf{Z}_n^{(1)}(\Phi;\theta^\star)) = \mathcal{O}(|\Lambda_n|^{-1})$. So

$$\begin{aligned}|\Lambda_n|^{1/2}\Big((\widehat{\theta}_n(\Phi)-\theta^\star) + \underline{\mathbf{H}}^{-1}(\theta^\star)\mathbf{Z}_n^{(1)}(\Phi;\theta^\star)\Big) &= -|\Lambda_n|^{1/2}\left(\underline{\mathbf{H}}_n^{-1}(\Phi;\widehat{\theta}_n,\theta^\star) - \underline{\mathbf{H}}^{-1}(\theta^\star)\right)\mathbf{Z}_n^{(1)}(\Phi;\theta^\star)\\ &\to 0,\end{aligned}$$

in probability as $n\to+\infty$. This implies (23). Finally, $\mathbf{U}_\Lambda(\varphi;\theta^\star)$ fulfills properties $(i)-(v)$ in **[E2(bis)]** for the same reasons as in the proof of Proposition 13 and, for (iii), from the proof of Theorem 2 (step 1, p. 257) in Billiot et al. (2008). ■

Remark 12. *In the same spirit as Remark 11, let us underline that the MPLE also satisfies **[E1]** and **[E2(bis)]** for some non locally stable and non exponential family models, including Lennard-Jones type models (provided a locality assumption).*

6.3 Estimation of asymptotic covariance matrices when considering the MPLE

We still focus on exponential family models. As in Section 4.3, we assume that the point process is observed in the domain $\Lambda_{n_0} \oplus D^+$ where $D^+ \geq D$ and Λ_{n_0} is a cube. Moreover, we consider the decomposition $\Lambda_{n_0} = \cup_{k \in \mathcal{K}_{n_0}} \Delta_k(\delta)$, for any δ such that $|\Lambda_{n_0}|\delta^{-d} \in \mathbb{N}$, where the $\Delta_k(\delta)$'s are disjoint cubes with side-length δ and centered, without loss of generality, at $k\delta$.

From (8) and (23), we have under the assumptions **[Exp]** and when considering the MPLE

$$R_{\infty,\Lambda}(\varphi; h, \theta^\star) := I_\Lambda(\varphi; h, \theta^\star) - \mathbf{LPL}^{(1)}(\varphi; \theta^\star)^T \mathbf{W}(h, \theta^\star) \tag{26}$$

where $\mathbf{W}(h,\theta) := \underline{\mathbf{H}}(\theta)^{-1}\mathcal{E}(h,\theta)$. A natural estimator of $\mathbf{W}(h,\theta^\star)$ is given by $\widehat{\mathbf{W}}_{n_0}(\varphi; h, \widehat{\theta}_{n_0}) := \widehat{\underline{\mathbf{H}}}_{n_0}(\varphi; \widehat{\theta}_{n_0})^{-1}\widehat{\mathcal{E}}_{n_0}(\varphi; h, \widehat{\theta}_{n_0})$ where

$$\begin{aligned}
\widehat{\underline{\mathbf{H}}}_{n_0}(\varphi; \widehat{\theta}_{n_0}) &= |\Lambda_{n_0}|^{-1} \int_{\Lambda_{n_0} \times \mathbb{M}} \mathbf{v}(x^m|\varphi)\mathbf{v}(x^m|\varphi)^T e^{-\widehat{\theta}_{n_0}^T \mathbf{v}(x^m|\varphi)} \mu(dx^m), \\
\widehat{\mathcal{E}}_{n_0}(\varphi; h, \widehat{\theta}_{n_0}) &= |\Lambda_{n_0}|^{-1} \int_{\Lambda_{n_0} \times \mathbb{M}} h(x^m, \varphi; \widehat{\theta}_{n_0})\mathbf{v}(x^m|\varphi) e^{-\widehat{\theta}_{n_0}^T \mathbf{v}(x^m|\varphi)} \mu(dx^m).
\end{aligned}$$

In this spirit, let $\widehat{R}_{n_0,\infty,\Lambda}(\varphi; h, \widehat{\theta}_{n_0}) := I_\Lambda(\varphi; h, \widehat{\theta}_{n_0}) - \mathbf{LPL}_\Lambda^{(1)}(\varphi; \widehat{\theta}_{n_0})^T \widehat{\mathbf{W}}_{n_0}(\varphi; h, \widehat{\theta}_{n_0})$ and $\widehat{\mathbf{R}}_{n_0,\infty,\Lambda}(\varphi; \mathbf{h}, \widehat{\theta}_{n_0}) := \left(\widehat{R}_{n_0,\infty,\Lambda}(\varphi; h_j, \widehat{\theta}_{n_0})\right)_{j=1,\dots,s}$. Based on these notation, we obtain the following estimations for $\lambda_{Inn}, \lambda_{Res}$ and $\underline{\mathbf{\Sigma}}_2(\theta^\star)$

$$\begin{aligned}
\widehat{\lambda}_{n_0,Inn}(\varphi, \widehat{\theta}_{n_0}(\varphi), \delta, D^\vee) &= |\Lambda_{n_0}|^{-1} \sum_{i \in \mathcal{K}_{n_0}} \sum_{j \in \mathbb{B}_i(\lceil \frac{D^\vee}{\delta} \rceil) \cap \mathcal{K}_{n_0}} I_{\Delta_i(\delta)}\left(\varphi; \widehat{\theta}_{n_0}(\varphi)\right) I_{\Delta_j(\delta)}\left(\varphi; \widehat{\theta}_{n_0}(\varphi)\right), \\
\widehat{\lambda}_{n_0,Res}(\varphi, \widehat{\theta}_{n_0}(\varphi), \delta, D^\vee) &= |\Lambda_{n_0}|^{-1} \sum_{i \in \mathcal{K}_{n_0}} \sum_{j \in \mathbb{B}_i(\lceil \frac{D^\vee}{\delta} \rceil) \cap \mathcal{K}_{n_0}} \widehat{R}_{\infty,\Delta_i(\delta)}(\varphi; h, \widehat{\theta}_{n_0}) \widehat{R}_{\infty,\Delta_j(\delta)}(\Phi; h, \widehat{\theta}_{n_0}), \\
\widehat{\underline{\mathbf{\Sigma}}}_{2n_0}(\varphi, \widehat{\theta}_{n_0}(\varphi), \delta, D^\vee) &= |\Lambda_{n_0}|^{-1} \sum_{i \in \mathcal{K}_{n_0}} \sum_{j \in \mathbb{B}_i(\lceil \frac{D^\vee}{\delta} \rceil) \cap \mathcal{K}_{n_0}} \widehat{\mathbf{R}}_{\infty,\Delta_i(\delta)}(\varphi; \mathbf{h}, \widehat{\theta}_{n_0}) \widehat{\mathbf{R}}_{\infty,\Delta_j(\delta)}(\varphi; \mathbf{h}, \widehat{\theta}_{n_0})^T.
\end{aligned}$$

Corollary 15. *Under the notation and assumptions of Propositions 4 and 6, and under **[Exp]**, then, for any $\delta > 0$ as above, one can consider a sequence δ_n which satisfies $\delta_{n_0} = \delta$ and $\delta_n \to \delta$, such that for any $D^\vee \geq D$, the estimators $\widehat{\lambda}_{n,Inn}(\Phi, \widehat{\theta}_n(\Phi), \delta_n, D^\vee)$, $\widehat{\lambda}_{n,Res}(\Phi, \widehat{\theta}_n(\Phi), \delta_n, D^\vee)$ and $\widehat{\underline{\mathbf{\Sigma}}}_{2n}(\Phi, \widehat{\theta}_n(\Phi), \delta_n, D^\vee)$ converge in probability (as $n \to +\infty$) towards respectively λ_{Inn}, λ_{Res} and $\underline{\mathbf{\Sigma}}_2(\theta^\star)$.*

Proof. We apply Proposition 9, where for any $\theta \in \mathcal{V}(\theta^\star)$, we set

- for λ_{Inn}: $\widehat{\mathbf{Y}}_\Lambda(\varphi;\theta) = \mathbf{Y}_\Lambda(\varphi;\theta) = I_\Lambda(\varphi; h, \theta)$.
- for λ_{Res}: $\mathbf{Y}_\Lambda(\varphi;\theta) = R_{\infty,\Lambda}(\varphi; h, \theta)$ and $\widehat{\mathbf{Y}}_{n,\Lambda}(\varphi;\theta) = \widehat{R}_{n,\infty,\Lambda}(\varphi; h, \theta)$.
- for $\underline{\mathbf{\Sigma}}_2(\theta^\star)$: $\mathbf{Y}_\Lambda(\varphi;\theta) = \mathbf{R}_{\infty,\Lambda}(\varphi; \mathbf{h}, \theta)$ and $\widehat{\mathbf{Y}}_{n,\Lambda}(\varphi;\theta) = \widehat{\mathbf{R}}_{n,\infty,\Lambda}(\varphi; \mathbf{h}, \theta)$.

The result is obvious for λ_{Inn}. For λ_{Res} (the proof is similar for $\underline{\mathbf{\Sigma}}_2(\theta^\star)$), it remains to prove that for any $\theta \in \mathbf{\Theta}$, $\sup_{k \in \mathcal{K}_n} \left| \widehat{R}_{n,\infty,\Delta_k(\delta_n)}(\Phi; h, \theta) - R_{n,\infty,\Delta_k(\delta_n)}(\Phi; h, \theta) \right| \to 0$ in probability as $n \to +\infty$. For any $k \in \mathcal{K}_n$, we derive

$$\widehat{R}_{n,\infty,\Delta_k(\delta_n)}(\Phi; h, \theta) - R_{\infty,\Delta_k(\delta_n)}(\Phi; h, \theta) = \mathbf{LPL}^{(1)}_{\Delta_k(\delta_n)}(\varphi; \theta)^T \left(\mathbf{W}(h,\theta) - \widehat{\mathbf{W}}_n(\varphi; h, \theta)\right).$$

Assumption **[E2(bis)]** (implied by **[Exp]**, see Proposition 14) ensures that $|\mathbf{LPL}^{(1)}_{\Delta_k(\delta_n)\setminus\Delta_k(\delta)}(\Phi;\theta)|$ converges to 0 in quadratic mean. In particular, the convergence of $\mathbf{LPL}^{(1)}_{\Delta_k(\delta_n)}(\Phi;\theta)$ towards $\mathbf{LPL}^{(1)}_{\Delta_k(\delta)}(\Phi;\theta)$ holds in probability. Moreover under the assumptions **[N1]** and **[E2(bis)]**, the ergodic theorem of Nguyen and Zessin (1979b) may be applied to prove that $\widehat{\mathbf{W}}_n(\Phi;h,\theta)$ converges almost surely towards $\mathbf{W}(h,\theta)$, as $n\to+\infty$. Slutsky's theorem ends the proof. ■

Remark 13. *If the model is not an exponential model, Corollary 15 still holds by replacing the vector of the sufficient statistics, $\mathbf{v}(x^m|\varphi)$, by the gradient vector of the local energy function, $\mathbf{V}^{(1)}(x^m|\varphi)$ in the different definitions.*

6.4 Positive definiteness of covariance matrices when considering the MPLE

Let us now focus on the positive-definitess of the above quantities. According to Proposition 7 the key assumption to check is **[PD]**.

As adressed in Remark 8, we begin by giving a general result ensuring that $\lambda_{Inn}>0$.

Proposition 16. *Under the assumption **[Exp]**, then $\lambda_{Inn}>0$ for the raw residuals, the Pearson residuals and the inverse residuals.*

Proof. In **[PD]**, we fix $\overline{\delta}=D$ and $B=\emptyset$. Let us write $\overline{\Omega}:=\overline{\Omega}_{\emptyset}$. Consider the following events for some $n\geq 1$

$$A_0=\left\{\varphi\in\overline{\Omega}:\varphi(\Delta_0(\overline{\delta}))=0\right\}\quad\text{ and }\quad A_n=\left\{\varphi\in\overline{\Omega}:\varphi(\Delta_0(\overline{\delta}))=n\right\},$$

and let $\varphi_0\in A_0$ and $\varphi_n\in A_n$. Recall that the local stability property (ensured by **[Exp]**) asserts that there exists $K\geq 0$ such that $V\left(x^m|\varphi;\theta^\star\right)\geq -K$ for any $x^m\in\mathbb{S}$ and any $\varphi\in\Omega$. Now, let us consider the three type of residuals.

Raw residuals $(h=1)$. From the local stability property

$$|I_{\overline{\Lambda}}(\varphi_n;h,\theta^\star)-I_{\overline{\Lambda}}(\varphi_0;h,\theta^\star)|\geq n-\left|\int_{\overline{\Lambda}\times\mathbb{M}}e^{-V(x^m|\varphi_n;\theta^\star)}-e^{-V(x^m|\varphi_0;\theta^\star)}\mu(dx^m)\right|\geq n-2|\overline{\Lambda}|e^K>0,$$

for n large enough. And so assuming that the left-hand-side is zero leads to a contradiction, which proves **[PD]**.

Inverse residuals $(h=e^V)$. Again, from the local stability property

$$|I_{\overline{\Lambda}}(\varphi_n;h,\theta^\star)-I_{\overline{\Lambda}}(\varphi_0;h,\theta^\star)|=\left|\sum_{x^m\in\varphi_{n\overline{\Lambda}}}e^{V(x^m|\varphi_n\setminus x^m;\theta^\star)}\right|\geq ne^{-K}>0,$$

which proves **[PD]** similarly to the previous case.

Pearson residuals $(h=e^{V/2})$. From the same argument

$$\begin{aligned}|I_{\overline{\Lambda}}(\varphi_n;h,\theta^\star)-I_{\overline{\Lambda}}(\varphi_0;h,\theta^\star)| \quad\geq\quad & \left|\sum_{x^m\in\varphi_{n\overline{\Lambda}}}e^{V(x^m|\varphi_n\setminus x^m;\theta^\star)/2}\right|-\\ & \left|\int_{\overline{\Lambda}\times\mathbb{M}}e^{-V(x^m|\varphi_n;\theta^\star)/2}-e^{-V(x^m|\varphi_0;\theta^\star)/2}\mu(dx^m)\right|\\ \geq\quad & ne^{-K/2}-2|\overline{\Lambda}|e^{K/2}>0,\end{aligned}$$

for n large enough, which ends the proof. ■

Proposition 16 asserts that **[PD]** is fullfilled for $\mathbf{Y}_{\overline{\Lambda}}(\Phi;\theta^\star)=I_{\overline{\Lambda}}(\Phi;\theta^\star)$. Therefore, the combination of Propositions 13, 14 and 16 and Corollary 15 ensures all the conditions of Corollary 10 hold. So a goodness-of-fit test based on (18) is valid for exponential family models satisfying **[Exp]** and for the raw residuals, the Pearson residuals and the inverse ones.

Now, let us focus on tests based on Corollary 11 and 12. The following result is important from a practical point of view. It asserts that λ_{Res} (and so $\underline{\mathbf{\Sigma}}_1(\theta^\star)$), and $\underline{\mathbf{\Sigma}}_2(\theta^\star)$ may fail to be positive-definite for an inappropriate choice of test function.

Proposition 17. *Let us consider an exponential family model, let $\widehat{\theta} := \widehat{\theta}^{MPLE}$ and let us choose a test function of the form $h(x^m, \varphi; \theta) = \omega^T \mathbf{v}(x^m|\varphi)$ for some $\omega \in \mathbb{R}^p \setminus 0$, then $\lambda_{Res} = 0$ and the matrices $\underline{\mathbf{\Sigma}}_1(\theta^\star)$ and $\underline{\mathbf{\Sigma}}_2(\theta^\star)$ in Propositions 4 and 6 are only semidefinite-positive matrices.*

Proof. The result is proved by noticing that

$$\begin{aligned}
\underline{\mathbf{H}}(\theta^\star)\,\omega &= \mathbf{E}\left(\mathbf{v}(0^M|\Phi)\mathbf{v}(0^M|\Phi)^T e^{-V\left(0^M|\Phi;\theta^\star\right)}\right)\omega \\
&= \mathbf{E}\left(\mathbf{v}(0^M|\Phi)\left(\omega^T\mathbf{v}(0^M|\Phi)\right)^T e^{-V\left(0^M|\Phi;\theta^\star\right)}\right) \\
&= \mathbf{E}\left(h(0^M,\Phi;\theta^\star)\mathbf{v}(0^M|\Phi)e^{-V\left(0^M|\Phi;\theta^\star\right)}\right) = \mathcal{E}(\omega^T\mathbf{v},\theta^\star).
\end{aligned}$$

Therefore, $\mathbf{W}\left(\omega^T\mathbf{v},\theta^\star\right) = \underline{\mathbf{H}}(\theta^\star)^{-1}\mathcal{E}(\omega^T\mathbf{v},\theta^\star) = \omega$, which means that for any $\varphi \in \Omega$ and any bounded domain Λ

$$R_{\infty,\Lambda}(\varphi;\omega^T\mathbf{v},\theta^\star) = I_\Lambda(\varphi;\omega^T\mathbf{v},\theta^\star) - \mathbf{LPL}_\Lambda^{(1)}(\varphi;\theta^\star)^T\omega = 0.$$

This means that if, for the framework 1, the test function is of the form $h = \omega^T\mathbf{v}$ then $\lambda_{Res} = 0$ and if one of the test functions, for the framework 2, is of the form $h = \omega^T\mathbf{v}$, then $\underline{\mathbf{\Sigma}}_2(\theta^\star)$ is necessary singular. ■

Remark 14. *As for Corollary 15, the result of Proposition 17 still holds in general by replacing the vector $\mathbf{v}(x^m|\varphi)$ by the gradient vector of the local energy function $\mathbf{V}^{(1)}(x^m|\varphi)$.*

As a consequence of Proposition 17, the two goodness-of-fit tests based on $T'_{1,n}$ and $T_{2,n}$ in Section 5.2 and 5.3 are not available (for the MPLE) if the test function h is a linear combination of the sufficient statistics $\mathbf{v}(x^m|\varphi)$. Since for most classical models, the value 1 can be obtained from a linear combination of $\mathbf{v}(x^m|\varphi)$, the raw residuals ($h = 1$) are not an appropriate choice for these two tests. This is the case for the two following examples: the area-interaction point process and the 2-type marked Strauss point process, which are presented in details in Appendix B. The following result proves that for a different choice of $h-$residuals, $\underline{\mathbf{\Sigma}}_1(\theta^\star)$ and $\underline{\mathbf{\Sigma}}_2(\theta^\star)$ are positive-definite.

Proposition 18. *For the 2-type marked Strauss point process and the area-interaction point process, when considering the MPLE as an estimator of $\theta^\star$, then*

- *the matrix $\underline{\mathbf{\Sigma}}_1(\theta^\star)$ obtained in Framework 1 from the inverse residuals $h = e^V$,*
- *the matrix $\underline{\mathbf{\Sigma}}_2(\theta^\star)$ obtained in Framework 2 from the empty space residuals, which are constructed for $0 < r_1 < \ldots < r_s < +\infty$ from the family of test functions*

$$h_j(x^m,\varphi;\theta) = \mathbf{1}_{[0,r_j]}(d(x^m,\varphi))e^{V(x^m|\varphi;\theta)}, \quad j = 1,\ldots,s,$$

are positive-definite.

The proof of this result is postponed in Appendix B. The combination of Propositions 13, 14, 18 and Corollary 15 ensures all the conditions of Corollary 11 and 12 hold. So a goodness-of-fit test based on (19) (resp. (20)) is valid for the 2-type marked Strauss point process and the area-interaction point process and for the inverse residuals (resp. the family of test functions based on the empty space function).

Following the Proof of Proposition 18, it is the belief of the authors that such a result holds for other models and other choices of test functions. However, another model and/or test functions will lead to a specific proof. Therefore, this result cannot be as general as the one presented in Proposition 16.

7 The non-hereditary case

Up to here, we have assumed through **[Mod-E]** that the family of energies is hereditary. We consider in this section the non-hereditary case. This particular situation can only occur in presence of a hardcore interaction. From a general point of view, we say that a family of energies involves a hardcore interaction if some point configurations have an infinite energy. Many classical models of Gibbs measures include a hardcore part, as the hard ball model.

A family of energies involving a hardcore part is hereditary if (2) holds. This is a common assumption done for Gibbs energies and it appears to be fulfilled in most classical models, including the hard ball model. However, one may encounter some non-hereditary models, in the sense that (2) does not hold. Intuitively, in this case, when one removes a point from an allowed point configuration, it is possible to obtain a forbidden point configuration. This occurs for instance for Gibbs Delaunay-Voronoï tessellations or forced-clustering processes (see Dereudre and Lavancier (2009) and Dereudre and Lavancier (2010)).

In the non-hereditary case, the GNZ formula (3), which is the basis to define the residuals, becomes false (see Remark 2 in Dereudre and Lavancier (2009)). It is extended to non-hereditary interactions in Dereudre and Lavancier (2009). This generalization requires to introduce the notion of removable points.

Definition 3. *Let $\varphi \in \Omega$ and $x \in \varphi$, then x is removable from φ if there exists $\Lambda \in \mathcal{B}(\mathbb{R}^d)$ such that $x \in \Lambda$ and $V_\Lambda(\varphi - x; \theta) < \infty$. The set of removable points in φ is denoted by $\mathcal{R}(\varphi)$.*

Notice that in the hereditary case, $\mathcal{R}(\varphi) = \varphi$.

The GNZ formula is then generalized to the non-hereditary case as follows. Assuming for any $\theta \in \boldsymbol{\Theta}$ that a Gibbs measure exists for the family of energies $(V_\Lambda(.;\theta))_{\Lambda\in\mathcal{B}(\mathbb{R}^d)}$, then, for any function $h(\cdot,\cdot,\theta) : \mathbb{S} \times \Omega \to \mathbb{R}^2$ such that the following quantities are finite,

$$\mathbf{E}\left(\int_{\mathbb{R}^d\times\mathbb{M}} h\left(x^m, \Phi; \theta\right) e^{-V(x^m|\Phi;\theta^\star)}\mu(dx^m)\right) = \mathbf{E}\left(\sum_{x^m\in\mathcal{R}(\Phi)} h\left(x^m, \Phi\setminus x^m; \theta\right)\right). \tag{27}$$

We can therefore define the $h-$residuals for (possibly) non-hereditary interactions. For any bounded domain Λ, if $\widehat{\theta}$ is an estimate of $\theta^\star$, the $h-$residuals are

$$R_\Lambda\left(\varphi; h, \widehat{\theta}\right) = \int_{\Lambda\times\mathbb{M}} h\left(x^m, \varphi; \widehat{\theta}\right) e^{-V\left(x^m|\varphi;\widehat{\theta}\right)}\mu(dx^m) - \sum_{x^m\in\mathcal{R}(\varphi_\Lambda)} h\left(x^m, \varphi\setminus x^m; \widehat{\theta}\right). \tag{28}$$

If the set of removable points $\mathcal{R}(\varphi)$ does not depend on θ, it is straightforward to extend all the asymptotic results obtained for the residuals in the preceding sections to (28).

If the set of removable points depends on θ, this is false. Even in the hereditary case, if θ is a hardcore parameter (as the hardcore distance in the hard ball model) then $\widehat{\theta}$ behaves as an estimator of the support of the distribution P_θ. In this case assumption **[E2]** has typically few chances to hold and the asymptotic law of the residuals is unknown. In Dereudre and Lavancier (2010) Figure 15, some simulations of raw-residuals for Gibbs Voronoï tessellations are presented, involving an estimated hardcore parameter in a non-hereditary setting : they show that the distribution of the residuals does not seem to be gaussian in this case.

8 Proofs

Since any stationary Gibbs measure can be represented as a mixture of ergodic measures (see Preston (1976)), it is sufficient to prove the different convergences involved in this paper for ergodic measures. We therefore assume from now on that $P_{\theta^\star}$ is ergodic.

8.1 Proof of Proposition 2

(a) Under **[C1]**, the ergodic theorem of Nguyen and Zessin (1979b) holds for both terms appearing in the definition of $I_{\tilde{\Lambda}_n}(\varphi; h, \theta^\star)$. Then, as $n \to +\infty$, one has $P_{\theta^\star}-$a.s.

$$|\tilde{\Lambda}_n|^{-1} I_{\tilde{\Lambda}_n}(\Phi; h, \theta^\star) \to \mathbf{E}\left(h\left(0^M, \Phi; \theta^\star\right) e^{-V\left(0^M|\Phi;\theta^\star\right)}\right) - \mathbf{E}\left(h\left(0^M, \Phi \setminus 0^M; \theta^\star\right)\right),$$

which equals to 0 from the GNZ formula (4).

(b) The aim is to prove that the difference $|\tilde{\Lambda}_n|^{-1} R_{\tilde{\Lambda}_n}\left(\varphi; h, \widehat{\theta}_n(\varphi)\right) - |\tilde{\Lambda}_n|^{-1} I_{\tilde{\Lambda}_n}(\varphi; h, \theta^\star)$ converges towards 0 for $P_{\theta^\star}-$a.e. φ. Let us write

$$R_{\tilde{\Lambda}_n}\left(\varphi; h, \widehat{\theta}_n(\varphi)\right) - I_{\tilde{\Lambda}_n}(\Phi; h, \theta^\star) := T_1(\varphi) - T_2(\varphi)$$

with

$$T_1(\varphi) := \int_{\tilde{\Lambda}_n \times \mathbb{M}} \left(f\left(x^m, \varphi; \widehat{\theta}_n(\varphi)\right) - f\left(x^m, \varphi; \theta^\star\right)\right) \mu(dx^m) \tag{29}$$

$$T_2(\varphi) := \sum_{x^m \in \varphi_{\tilde{\Lambda}_n}} h\left(x^m, \varphi \setminus x^m; \widehat{\theta}_n(\varphi)\right) - h\left(x^m, \varphi \setminus x^m; \theta^\star\right). \tag{30}$$

Under the Assumptions **[C2]** and **[E1]**, from the ergodic theorem and the GNZ formula, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$

$$\begin{aligned}
|\tilde{\Lambda}_n|^{-1} T_1(\varphi) &\leq \frac{2}{|\tilde{\Lambda}_n|} \int_{\tilde{\Lambda}_n \times \mathbb{M}} \left(\widehat{\theta}_n(\varphi) - \theta^\star\right)^T \mathbf{f}^{(1)}\left(x^m, \varphi; \theta^\star\right) \mu(dx^m) \\
&\leq 2\|\widehat{\theta}_n(\varphi) - \theta^\star\| \times \frac{1}{|\tilde{\Lambda}_n|} \int_{\tilde{\Lambda}_n \times \mathbb{M}} \|\mathbf{f}^{(1)}\left(x^m, \varphi; \theta^\star\right)\| \mu(dx^m) \\
&\leq 4\|\widehat{\theta}_n(\varphi) - \theta^\star\| \times \mathbf{E}\left(\|\mathbf{f}^{(1)}\left(0^M, \Phi; \theta^\star\right)\|\right),
\end{aligned} \tag{31}$$

and

$$\begin{aligned}
|\tilde{\Lambda}_n|^{-1} T_2(\varphi) &\leq \frac{2}{|\tilde{\Lambda}_n|} \sum_{x^m \in \varphi_{\tilde{\Lambda}_n}} \left(\widehat{\theta}_n(\varphi) - \theta^\star\right)^T \mathbf{h}^{(1)}\left(x^m, \varphi \setminus x^m; \theta^\star\right) \\
&\leq 4\|\widehat{\theta}_n(\varphi) - \theta^\star\| \times \mathbf{E}\left(\|\mathbf{h}^{(1)}\left(0^M, \Phi; \theta^\star\right)\| e^{-V\left(0^M|\Phi;\theta^\star\right)}\right).
\end{aligned} \tag{32}$$

Equations (31) and (32) lead to

$$|\tilde{\Lambda}_n|^{-1} R_{\tilde{\Lambda}_n}\left(\varphi; h, \widehat{\theta}_n(\varphi)\right) - |\tilde{\Lambda}_n|^{-1} I_{\tilde{\Lambda}_n}(\varphi; h, \theta^\star) \leq c\|\widehat{\theta}_n(\varphi) - \theta^\star\|,$$

for n large enough, with $c = 4 \times \mathbf{E}\left(\|\mathbf{f}^{(1)}\left(0^M, \Phi; \theta^\star\right)\| + \|\mathbf{h}^{(1)}\left(0^M, \Phi; \theta^\star\right)\| e^{-V\left(0^M|\Phi;\theta^\star\right)}\right)$.

8.2 Proof of Proposition 3

Recall that

$$R_{\tilde{\Lambda}_n}\left(\varphi; h, \widehat{\theta}_n(\varphi)\right) - I_{\tilde{\Lambda}_n}(\varphi; h, \theta^\star) = T_1(\varphi) - T_2(\varphi)$$

where $T_1(\varphi)$ and $T_2(\varphi)$ are defined by (29) and (30). Let us write

$$\begin{aligned}
T_1(\varphi) &= \int_{\tilde{\Lambda}_n \times \mathbb{M}} \left(\widehat{\theta}_n(\varphi) - \theta^\star\right)^T \mathbf{f}^{(1)}\left(x^m, \varphi; \theta^\star\right) \mu(dx^m) + T_1'(\varphi) \\
T_2(\varphi) &= \sum_{x^m \in \varphi_{\tilde{\Lambda}_n}} \left(\widehat{\theta}_n(\varphi) - \theta^\star\right)^T \mathbf{h}^{(1)}\left(x^m, \varphi \setminus x^m; \theta^\star\right) + T_2'(\varphi),
\end{aligned}$$

with

$$\begin{aligned} T_1'(\varphi) &:= \int_{\tilde{\Lambda}_n\times\mathbb{M}} A_1\left(x^m,\varphi;\widehat{\theta}_n(\varphi)\right)\mu(dx^m) \\ T_2'(\varphi) &= \sum_{x^m\in\varphi_{\tilde{\Lambda}_n}} A_2\left(x^m,\varphi\setminus x^m;\widehat{\theta}_n(\varphi)\right) \end{aligned}$$

and

$$\begin{aligned} A_1\left(x^m,\varphi;\widehat{\theta}_n(\varphi)\right) &:= f\left(x^m,\varphi;\widehat{\theta}_n\right) - f\left(x^m,\varphi;\theta^\star\right) - \left(\widehat{\theta}_n(\varphi)-\theta^\star\right)^T \mathbf{f}^{(1)}\left(x^m,\varphi;\theta^\star\right) \\ A_2\left(x^m,\varphi;\widehat{\theta}_n(\varphi)\right) &:= h\left(x^m,\varphi;\widehat{\theta}_n\right) - h\left(x^m,\varphi;\theta^\star\right) - \left(\widehat{\theta}_n(\varphi)-\theta^\star\right)^T \mathbf{h}^{(1)}\left(x^m,\varphi;\theta^\star\right). \end{aligned}$$

From the mean value theorem, there exist for $j=1,\dots,p$,
$\xi_{1,j},\xi_{2,j}\in[\min(\widehat{\theta}_1,\theta_1^\star),\max(\widehat{\theta}_1,\theta_1^\star)]\times\ldots\times[\min(\widehat{\theta}_p,\theta_p^\star),\max(\widehat{\theta}_p,\theta_p^\star)]$ such that

$$A_1\left(x^m,\varphi;\widehat{\theta}_n(\varphi)\right) = \sum_{j=1}^p(\widehat{\theta}_j-\theta_j^\star)\left(f_j^{(1)}\left(x^m,\varphi;\xi_{1,j}\right)-f_j^{(1)}\left(x^m,\varphi;\theta^\star\right)\right) \quad (33)$$

$$A_2\left(x^m,\varphi\setminus x^m;\widehat{\theta}_n(\varphi)\right) = \sum_{j=1}^p(\widehat{\theta}_j-\theta_j^\star)\left(h_j^{(1)}\left(x^m,\varphi\setminus x^m;\xi_{2,j}\right)-h_j^{(1)}\left(x^m,\varphi\setminus x^m;\theta^\star\right)\right). \quad (34)$$

Let $j\in\{1,\dots,p\}$, again from the mean value theorem, there exist for $\ell=1,2$ and for $k=1,\dots,p$, $\eta_{\ell,j,k}\in[\min(\xi_{\ell,j,1},\theta_1^\star),\max(\xi_{\ell,j,1},\theta_1^\star)]\times\ldots\times[\min(\xi_{\ell,j,p},\theta_p^\star),\max(\xi_{\ell,j,p},\theta_p^\star)]$ such that

$$f_j^{(1)}\left(x^m,\varphi;\xi_{1,j}\right)-f_j^{(1)}\left(x^m,\varphi;\theta^\star\right) = \sum_{k=1}^p(\xi_{1,j,k}-\theta_k^\star)\,f_{jk}^{(2)}\left(x^m,\varphi;\eta_{1,j,k}\right) \quad (35)$$

$$h_j^{(1)}\left(x^m,\varphi\setminus x^m;\xi_{2,j}\right)-h_j^{(1)}\left(x^m,\varphi\setminus x^m;\theta^\star\right) = \sum_{k=1}^p(\xi_{2,j,k}-\theta_k^\star)\,h_{jk}^{(2)}\left(x^m,\varphi\setminus x^m;\eta_{2,j,k}\right) \quad (36)$$

By combining (33), (34), (35) and (36) and under **[N1]**, we can deduce the existence of $n_0\in\mathbb{N}$ such that for all $n\geq n_0$, one has for $P_{\theta^\star}-$a.e. φ

$$\begin{aligned} |\tilde{\Lambda}_n|^{-1}|T_1'(\varphi)| &\leq \frac{2}{|\tilde{\Lambda}_n|}\int_{\tilde{\Lambda}_n\times\mathbb{M}}\sum_{j,k}\left|(\widehat{\theta}_j-\theta_j^\star)(\widehat{\theta}_k-\theta_k^\star)f_{jk}^{(2)}\left(x^m,\varphi;\theta^\star\right)\right|\mu(dx^m) \\ &\leq 2\|\widehat{\theta}_n(\varphi)-\theta^\star\|^2\times\frac{1}{|\tilde{\Lambda}_n|}\int_{\tilde{\Lambda}_n\times\mathbb{M}}\|\underline{\mathbf{f}}^{(2)}\left(x^m,\varphi;\theta^\star\right)\|\mu(dx^m) \\ &\leq 4\|\widehat{\theta}_n(\varphi)-\theta^\star\|^2\times\mathbf{E}\left(\|\underline{\mathbf{f}}^{(2)}\left(0^M,\Phi;\theta^\star\right)\|\right) \end{aligned}$$

and

$$\begin{aligned} |\tilde{\Lambda}_n|^{-1}|T_2'(\varphi)| &\leq \frac{2}{|\tilde{\Lambda}_n|}\sum_{x^m\in\varphi_{\tilde{\Lambda}_n}}\sum_{j,k}\left|(\widehat{\theta}_j-\theta_j^\star)(\widehat{\theta}_k-\theta_k^\star)h_{jk}^{(2)}\left(x^m,\varphi\setminus x^m;\theta^\star\right)\right| \\ &\leq 2\|\widehat{\theta}_n(\varphi)-\theta^\star\|^2\times\frac{1}{|\tilde{\Lambda}_n|}\sum_{x^m\in\varphi_{\tilde{\Lambda}_n}}\|\underline{\mathbf{h}}^{(2)}\left(x^m,\varphi\setminus x^m;\theta^\star\right)\| \\ &\leq 4\|\widehat{\theta}_n(\varphi)-\theta^\star\|^2\times\mathbf{E}\left(\|\underline{\mathbf{h}}^{(2)}\left(0^M,\Phi;\theta^\star\right)\|e^{-V\left(0^M|\Phi;\theta^\star\right)}\right) \end{aligned}$$

Since

$$|\tilde{\Lambda}_n|^{1/2}\|\widehat{\theta}_n(\varphi)-\theta^\star\|^2 = \left(\frac{|\tilde{\Lambda}_n|}{|\Lambda_n|}\right)^{1/2}\|\,|\Lambda_n|^{1/2}(\widehat{\theta}_n(\varphi)-\theta^\star)\|\times\|\widehat{\theta}_n(\varphi)-\theta^\star\|$$

then, under the assumptions of Proposition 3, one has, from Slustsky's theorem, the following convergence in probability as $n \to +\infty$

$$|\tilde{\Lambda}_n|^{1/2} \|\widehat{\theta}_n(\Phi) - \theta^\star\|^2 \xrightarrow{P} 0.$$

By combining all these results, one obtains the following convergence in probability, as $n \to +\infty$

$$|\tilde{\Lambda}_n|^{-1/2} \left(T_1(\Phi) - T_2(\Phi) - |\tilde{\Lambda}_n| \left(\widehat{\theta}_n(\Phi) - \theta^\star \right)^T \mathbf{X}_{\tilde{\Lambda}_n}(\Phi) \right) = |\tilde{\Lambda}_n|^{-1/2} \left(T_1'(\Phi) - T_2'(\Phi) \right) \xrightarrow{P} 0.$$

where $\mathbf{X}_{\tilde{\Lambda}_n}(\Phi)$ is the random vector defined for all $j = 1, \ldots, p$ by

$$\left(\mathbf{X}_{\tilde{\Lambda}_n}(\Phi) \right)_j := \frac{1}{|\tilde{\Lambda}_n|} \int_{\tilde{\Lambda}_n \times \mathbb{M}} f_j^{(1)}(x^m, \Phi; \theta^\star) \, \mu(dx^m) - \frac{1}{|\tilde{\Lambda}_n|} \sum_{x^m \in \Phi_{\tilde{\Lambda}_n}} h_j^{(1)}(x^m, \Phi \setminus x^m; \theta^\star).$$

By using the ergodic theorem and the GNZ formula, one has $P_{\theta^\star}-$a.s. as $n \to +\infty$

$$\left(\mathbf{X}_{\tilde{\Lambda}_n}(\Phi) \right)_j \to \mathbf{E} \left(f_j^{(1)}(0^M, \Phi; \theta^\star) - h_j^{(1)}(0^M, \Phi; \theta^\star) \, e^{-V(0^M | \Phi; \theta^\star)} \right).$$

Finally, let us notice that for all $(m, \varphi) \in \mathbb{M} \times \Omega$ and for all $j = 1, \ldots, p$

$$\begin{aligned} f_j^{(1)}(0^m, \varphi; \theta^\star) &= \frac{\partial}{\partial \theta_j} \left(h(0^m, \varphi; \theta) \, e^{-V(0^m | \varphi; \theta)} \right) \Big|_{\theta = \theta^\star} \\ &= h_j^{(1)}(0^m, \varphi; \theta^\star) \, e^{-V(0^m | \varphi; \theta^\star)} - h(0^m, \varphi; \theta) \, V_j^{(1)}(0^m | \varphi; \theta^\star) \, e^{-V(0^m | \varphi; \theta^\star)}. \end{aligned}$$

Therefore $\left(\mathbf{X}_{\tilde{\Lambda}_n}(\Phi) \right)_j \to -\mathcal{E}_j(h, \theta^\star)$ $P_{\theta^\star}-$a.s. as $n \to +\infty$. This finally leads to the following convergence in probability, as $n \to +\infty$

$$|\tilde{\Lambda}_n|^{-1/2} \left(T_1(\Phi) - T_2(\Phi) + |\tilde{\Lambda}_n| \left(\widehat{\theta}_n(\Phi) - \theta^\star \right)^T \mathcal{E}(h; \theta^\star)) \right) \xrightarrow{P} 0.$$

8.3 Proof of Proposition 4

Let us first state a result widely used in the following.

Lemma 19. *For any bounded domain Λ and for any test function h*

$$\mathbf{E} \left(I_\Lambda(\Phi; h, \theta^\star) | \Phi_{\Lambda^c} \right) = 0. \tag{37}$$

The proof of Lemma 19 is omitted since it corresponds to the proof of Theorem 2 (Step 1, p. 257) of Billiot et al. (2008) by subsituting $v_j(x^m|\varphi)$ by the test function $h(x^m, \varphi; \theta^\star)$.

For all $n \in \mathbb{N}$, the domain Λ_n is assumed to be a cube divided as $\Lambda_n = \bigcup_{j \in \mathcal{J}} \Lambda_{j,n}$ where for all $j \in \mathcal{J}$, the $\Lambda_{j,n}$'s are disjoint cubes. So $|\Lambda_n| = |\mathcal{J}||\Lambda_{j,n}| = |\mathcal{J}||\Lambda_{0,n}|$. Moreover, for all $j \in \mathcal{J}$, we can decompose each $\Lambda_{j,n}$ in the following way :

$$\Lambda_{j,n} := \bigcup_{k \in \mathcal{K}_{j,n}} \Delta_k(D_n) \tag{38}$$

where the $\Delta_k(D_n)$'s are disjoint cubes with side-length D_n and $\mathcal{K}_{j,n} \subset \mathbb{Z}^d$. The side-length D_n is chosen greater than D and as close as possible to D, leading to

$$D_n = \frac{|\Lambda_n|^{1/d}}{|\mathcal{J}|^{1/d} \left\lfloor \frac{|\Lambda_n|^{1/d}}{|\mathcal{J}|^{1/d} D} \right\rfloor}.$$

This choice implies $D_n \to D$ when $n \to \infty$ and guarantees $D \le D_n \le 2D$ as soon as $|\Lambda_n| \ge |\mathcal{J}|D^d$. The cubes $\Lambda_{j,n}$'s are therefore divided into $|\mathcal{K}_{j,n}| = |\Lambda_{0,n}|D_n^{-d}$ cubes whose volumes are closed to D^d. Denoting $\mathcal{K}_n = \bigcup_{j\in\mathcal{J}} \mathcal{K}_{j,n}$, we have $|\mathcal{K}_n| = |\Lambda_n|D_n^{-d} = |\mathcal{J}||\mathcal{K}_{j,n}|$ and finally

$$\Lambda_n = \bigcup_{j\in\mathcal{J}} \bigcup_{k\in\mathcal{K}_{j,n}} \Delta_k(D_n) = \bigcup_{k\in\mathcal{K}_n} \Delta_k(D_n). \tag{39}$$

From Proposition 3 and under Assumption **[E2(bis)]**, one has for any $j \in \mathcal{J}$

$$|\Lambda_{j,n}|^{-1/2} R_{\Lambda_{j,n}}\left(\Phi; h, \widehat{\theta}_n(\Phi)\right) = |\Lambda_{j,n}|^{-1/2} R_{\infty,\Lambda_{j,n}}\left(\Phi; h, \theta^\star\right) + o_P(1),$$

where $R_{\infty,\Lambda_{j,n}}\left(\Phi; h, \theta^\star\right)$ is defined in (8).

Therefore the proof of Proposition 4 reduces to the proof of the asymptotic normality of the vector $\left(|\Lambda_{j,n}|^{-1/2} R_{\infty,\Lambda_{j,n}}\left(\Phi; h, \theta^\star\right)\right)_{j\in\mathcal{J}}$. Now

$$\begin{aligned}
|\Lambda_{j,n}|^{-1/2} R_{\infty,\Lambda_{j,n}}\left(\Phi; h, \theta^\star\right) &= |\Lambda_{0,n}|^{-1/2}\left(I_{\Lambda_{j,n}}\left(\Phi; h, \theta^\star\right) - \frac{|\Lambda_{0,n}|}{|\Lambda_n|}\mathbf{U}_{\Lambda_n}\left(\Phi; \theta^\star\right)^T \mathcal{E}(h;\theta^\star)\right)\\
&= \frac{|\Lambda_{0,n}|^{1/2}}{|\Lambda_n|}\left(|\mathcal{J}| \times I_{\Lambda_{j,n}}\left(\Phi; h, \theta^\star\right) - \mathbf{U}_{\Lambda_n}\left(\Phi; \theta^\star\right)^T \mathcal{E}(h;\theta^\star)\right)\\
&= \frac{1}{D_n^{d/2}|\mathcal{J}|^{1/2}} \frac{1}{|\mathcal{K}_n|^{1/2}} \sum_{k\in\mathcal{K}_n} W_{j,n,\Delta_k(D_n)}\left(\Phi; \theta^\star\right),
\end{aligned} \tag{40}$$

where for any $\varphi \in \Omega$

$$W_{j,n,\Delta_k(D_n)}\left(\varphi; \theta^\star\right) = \begin{cases} W^{(1)}_{\Delta_k(D_n)}\left(\varphi; \theta^\star\right) := |\mathcal{J}| \times I_{\Delta_k(D_n)}\left(\varphi; h, \theta^\star\right) & \\ \qquad\qquad -\mathbf{U}_{\Delta_k(D_n)}\left(\varphi; \theta^\star\right)^T \mathcal{E}(h;\theta^\star) & \text{if } k \in \mathcal{K}_{j,n}, \\ W^{(2)}_{\Delta_k(D_n)}\left(\varphi; \theta^\star\right) := -\mathbf{U}_{\Delta_k(D_n)}\left(\varphi; \theta^\star\right)^T \mathcal{E}(h;\theta^\star) & \text{if } k \in \mathcal{K}_n \setminus \mathcal{K}_{j,n}. \end{cases} \tag{41}$$

Therefore, to prove a central limit theorem for the vector $\left(|\Lambda_{j,n}|^{-1/2} R_{\infty,\Lambda_{j,n}}\left(\Phi; h, \theta^\star\right)\right)_{j\in\mathcal{J}}$, it suffices to apply Theorem 21 (see Appendix A), where in its statement we choose $\mathbf{Z}_{n,k} = (W_{j,n,\Delta_k(D_n)}(\Phi;\theta^\star))_{j\in\mathcal{J}}$, $X_{n,i} = \Phi_{\Delta_i(D_n)}$ and $p = |\mathcal{J}|$. For this, we first have to specify the asymptotic variance matrix $\underline{\mathbf{\Sigma}}$, then to check the assumptions of Theorem 21.

First step: computation of the asymptotic variance.

Let us fix a cartesian coordinate system such that 0 is the center of Λ_n. We assume, without loss of generality, that $|\mathcal{J}|$ is odd. Moreover, we can always choose an odd number $|\mathcal{K}_{j,n}|$ of cubes $\Delta_k(D_n)$ in (55). Consequently, $\Lambda_{0,n}$ may be centered at 0 and each $\Delta_k(D_n)$ is centered at kD_n, $k \in \mathbb{Z}^d$. Note that if $|\mathcal{J}|$ was even, each $\Delta_k(D_n)$ would be centered at $kD_n/2$. So, in this system, $\mathcal{K}_n$ is a subset of $\mathbb{Z}^d$, independent of D_n, with $|\mathcal{K}_n| = |\mathcal{J}|\left\lfloor \frac{|\Lambda_n|^{1/d}}{|\mathcal{J}|^{1/d}D} \right\rfloor^d$ elements.

Set, for all $k, k' \in \mathbb{Z}^d$,

$$\begin{cases} E^{(1)}_{k,k'}(D_n) := \mathbf{E}\left(W^{(1)}_{\Delta_k(D_n)}\left(\Phi; \theta^\star\right) W^{(1)}_{\Delta_{k'}(D_n)}\left(\Phi; \theta^\star\right)\right) \\ E^{(12)}_{k,k'}(D_n) := \mathbf{E}\left(W^{(1)}_{\Delta_k(D_n)}\left(\Phi; \theta^\star\right) W^{(2)}_{\Delta_{k'}(D_n)}\left(\Phi; \theta^\star\right)\right) \\ E^{(2)}_{k,k'}(D_n) := \mathbf{E}\left(W^{(2)}_{\Delta_k(D_n)}\left(\Phi; \theta^\star\right) W^{(2)}_{\Delta_{k'}(D_n)}\left(\Phi; \theta^\star\right)\right) \end{cases}$$

Note that from the stationarity of the point process, we have $E^{(l)}_{k,k'}(D_n) = E^{(l)}_{0,k-k'}(D_n)$, for $l = 1, 12, 2$. Moreover, under Assumptions **[N4]** and **[E2(bis)]**, for any $k \in \mathcal{K}_n$ and for any configuration φ, since $D_n \ge D$, $W^{(i)}_{\Delta_k(D_n)}\left(\varphi; \theta^\star\right)$, $i = 1, 2$, depends only on $\varphi_{\Delta_l(D_n)}$ for $|l - k| \le 1$

that is $l \in \mathbb{B}_k(1)$. As a consequence, if $k' \in \mathbb{B}_k^c(1)$, $W^{(i)}_{\Delta_{k'}(D_n)}(\Phi;\theta^\star)$ is a measurable function of $\Phi_{\Delta_k^c(D_n)}$. This leads, for $i,j=1,2$, to

$$\begin{aligned}
\mathbf{E}\left(W^{(i)}_{\Delta_k(D_n)}(\Phi;\theta^\star)\,W^{(j)}_{\Delta_{k'}(D_n)}(\Phi;\theta^\star)\right) &= \mathbf{E}\left(\mathbf{E}\left(W^{(i)}_{\Delta_k(D_n)}(\Phi;\theta^\star)\,W^{(j)}_{\Delta_{k'}(D_n)}(\Phi;\theta^\star)\,|\Phi_{\Delta_k^c(D_n)}\right)\right)\\
&= \mathbf{E}\left(W^{(i)}_{\Delta_{k'}(D_n)}(\Phi;\theta^\star)\,\mathbf{E}\left(W^{(j)}_{\Delta_k(D_n)}(\Phi;\theta^\star)\,|\Phi_{\Delta_k^c(D_n)}\right)\right).
\end{aligned} \tag{42}$$

From Lemma 19 and under **[E2(bis)]** then for any $k \in \mathbb{Z}^d$ and for $i=1,2$,

$$\mathbf{E}\left(W^{(i)}_{\Delta_k(D_n)}(\Phi;\theta^\star)\,|\Phi_{\Delta_k^c(D_n)}\right) = 0. \tag{43}$$

From (42) and (43), we deduce that, for $l=1,12,2$,

$$k' \in \mathbb{B}_k^c(1) \Longrightarrow E^{(l)}_{k,k'}(D_n) = 0. \tag{44}$$

We are now in position to compute the covariance. For any i and j in $\mathcal{J}$, from (40),

$$\begin{aligned}
&cov\left(|\Lambda_{i,n}|^{-1/2}R_{\infty,\Lambda_{i,n}}(\Phi;h,\theta^\star),|\Lambda_{j,n}|^{-1/2}R_{\infty,\Lambda_{j,n}}(\Phi;h,\theta^\star)\right)\\
&\qquad = \frac{1}{D_n^d|\mathcal{J}|}\mathbf{E}\left(\frac{1}{|\mathcal{K}_n|}\sum_{k\in\mathcal{K}_n}\sum_{k'\in\mathcal{K}_n} W_{i,n,\Delta_k(D_n)}(\Phi;\theta^\star)\,W_{j,n,\Delta_{k'}(D_n)}(\Phi;\theta^\star)\right).
\end{aligned} \tag{45}$$

Let us first consider the case $i=j$. We may write

$$\begin{aligned}
&\mathbf{E}\left(\frac{1}{|\mathcal{K}_n|}\sum_{k\in\mathcal{K}_n}\sum_{k'\in\mathcal{K}_n} W_{i,n,\Delta_k(D_n)}(\Phi;\theta^\star)^2\right)\\
&\quad = \frac{1}{|\mathcal{K}_n|}\Bigg(\underbrace{\sum_{k,k'\in\mathcal{K}_{i,n}} E^{(1)}_{k,k'}(D_n)}_{:=S_1} + 2\underbrace{\sum_{k\in\mathcal{K}_{i,n},k'\in\mathcal{K}_n\setminus\mathcal{K}_{i,n}} E^{(12)}_{k,k'}(D_n)}_{:=S_2} + \underbrace{\sum_{k,k'\in\mathcal{K}_n\setminus\mathcal{K}_{i,n}} E^{(2)}_{k,k'}(D_n)}_{:=S_3}\Bigg).
\end{aligned}$$

The following lemma will be useful to drop the dependence on D_n in each term S_1, S_2, S_3 above.

Lemma 20. *For any $i,j=1,2$, denoting $\overline{\Delta}_0(\tau) = \cup_{k\in\mathbb{B}_0(1)}\Delta_k(\tau)$ (for some $\tau>0$), we have*

$$W^{(i)}_{\overline{\Delta}_0(D_n)}(\Phi;\theta^\star)\,W^{(j)}_{\overline{\Delta}_0(D_n)}(\Phi;\theta^\star) \xrightarrow{L_1} W^{(i)}_{\overline{\Delta}_0(D)}(\Phi;\theta^\star)\,W^{(j)}_{\overline{\Delta}_0(D)}(\Phi;\theta^\star).$$

Proof. For any $i=1,2$, $W^{(i)}_{\overline{\Delta}_0(D_n)}$ is a linear combination of $I_{\overline{\Delta}_0(D_n)}$ and $\mathbf{U}_{\overline{\Delta}_0(D_n)}$, which converge respectively in L^2 to $I_{\overline{\Delta}_0(D)}$ and $\mathbf{U}_{\overline{\Delta}_0(D)}$ by **[N3]** and **[E2(bis)]**, since $D_n \to D$. Thus $W^{(i)}_{\overline{\Delta}_0(D_n)}$ converges in L^2 to $W^{(i)}_{\overline{\Delta}_0(D)}$ as $n\to\infty$. Similarly, for any $j=1,2$, $W^{(j)}_{\overline{\Delta}_0(D_n)}$ tends in L^2 to $W^{(j)}_{\overline{\Delta}_0(D)}$. The convergence stated in Lemma 20 then follows. ■

Let us focus on the asymptotic of each term S_1, S_2, S_3.

Term S_1: from (44),

$$S_1 = \sum_{k\in\mathcal{K}_{i,n}}\Bigg(\sum_{k'\in\mathbb{B}_k(1)\cap\mathcal{K}_{i,n}} E^{(1)}_{k,k'}(D_n) + \underbrace{\sum_{k'\in\mathbb{B}_k^c(1)\cap\mathcal{K}_{i,n}} E^{(1)}_{k,k'}(D_n)}_{=0}\Bigg) = \sum_{k\in\mathcal{K}_{i,n}}\sum_{k'\in\mathbb{B}_k(1)\cap\mathcal{K}_{i,n}} E^{(1)}_{k,k'}(D_n).$$

Let $\widetilde{\mathcal{K}}_{i,n} := \mathcal{K}_{i,n} \cap \left(\cup_{j\in\partial\mathcal{K}_{i,n}} \mathbb{B}_j(1)\right)$ and note that $\frac{|\widetilde{\mathcal{K}}_{i,n}|}{|\mathcal{K}_{i,n}|} \to 0$ as $n \to +\infty$. Then,

$$S_1 = \sum_{k\in\mathcal{K}_{i,n}\setminus\widetilde{\mathcal{K}}_{i,n}} \sum_{k'\in\mathbb{B}_k(1)\cap\mathcal{K}_{i,n}} E^{(1)}_{k,k'}(D_n) + \underbrace{\sum_{k\in\widetilde{\mathcal{K}}_{i,n}} \sum_{k'\in\mathbb{B}_k(1)\cap\mathcal{K}_{i,n}} E^{(1)}_{k,k'}(D_n)}_{:=A_1}.$$

Since,

$$\frac{1}{|\mathcal{K}_n|} \times |A_1| \leq \frac{|\widetilde{\mathcal{K}}_{i,n}|}{|\mathcal{K}_n|} \sum_{k\in\mathbb{B}_0(1)} |E^{(1)}_{0,k}(D_n)| \overset{n\to+\infty}{\longrightarrow} 0,$$

(because $D \leq D_n \leq 2D$ and $\frac{|\widetilde{\mathcal{K}}_{i,n}|}{|\mathcal{K}_{i,n}|} \to 0$), we obtain, as $n \to +\infty$,

$$\frac{1}{|\mathcal{K}_n|}\, S_1 \;\sim\; \frac{|\mathcal{K}_{i,n}\setminus\widetilde{\mathcal{K}}_{i,n}|}{|\mathcal{K}_n|} \sum_{k\in\mathbb{B}_0(1)} E^{(1)}_{0,k}(D_n) \sim \frac{|\mathcal{K}_{i,n}|}{|\mathcal{K}_n|} \sum_{k\in\mathbb{B}_0(1)} E^{(1)}_{0,k}(D_n).$$

From Lemma 20,

$$\sum_{k\in\mathbb{B}_0(1)} E^{(1)}_{0,k}(D_n) = \mathbf{E}\left(W^{(1)}_{\Delta_0(D_n)}(\Phi;\theta^\star)\, W^{(1)}_{\Delta_0(D_n)}(\Phi;\theta^\star)\right) \longrightarrow \sum_{k\in\mathbb{B}_0(1)} E^{(1)}_{0,k}(D).$$

Therefore,

$$\frac{1}{|\mathcal{K}_n|}\, S_1 \sim \frac{1}{|\mathcal{J}|} \sum_{k\in\mathbb{B}_0(1)} E^{(1)}_{0,k}(D).$$

Term S_2: with similar arguments as above, we obtain

$$\begin{aligned} S_2 \;&=\; \underbrace{\sum_{k\in\mathcal{K}_{i,n}\setminus\widetilde{\mathcal{K}}_{i,n}} \sum_{k'\in\mathcal{K}_n\setminus\mathcal{K}_{i,n}} E^{(12)}_{k,k'}(D_n)}_{=0} \\ &\quad + \sum_{k\in\widetilde{\mathcal{K}}_{i,n}} \left(\sum_{k'\in\mathbb{B}_k(1)\cap(\mathcal{K}_n\setminus\mathcal{K}_{i,n})} E^{(12)}_{k,k'}(D_n) + \underbrace{\sum_{k'\in\mathbb{B}^c_k(1)\cap(\mathcal{K}_n\setminus\mathcal{K}_{i,n})} E^{(12)}_{k,k'}(D_n)}_{=0} \right) \end{aligned}$$

Therefore, since $\frac{|\widetilde{\mathcal{K}}_{i,n}|}{|\mathcal{K}_n|} \to 0$ and $D \leq D_n \leq 2D$,

$$\frac{1}{|\mathcal{K}_n|}\, S_2 \leq \frac{|\widetilde{\mathcal{K}}_{i,n}|}{|\mathcal{K}_n|} \sum_{k\in\mathbb{B}_0(1)} |E^{(12)}_{0,k}(D_n)| \overset{n\to+\infty}{\longrightarrow} 0.$$

Term S_3:

$$S_3 = \sum_{k\in\mathcal{K}_n\setminus\mathcal{K}_{i,n}} \sum_{k'\in\mathbb{B}_k(1)\cap(\mathcal{K}_n\setminus\mathcal{K}_{i,n})} E^{(2)}_{k,k'}(D_n) + \underbrace{\sum_{k\in\mathcal{K}_n\setminus\mathcal{K}_{i,n}} \sum_{k'\in\mathbb{B}^c_k(1)\cap(\mathcal{K}_n\setminus\mathcal{K}_{i,n})} E^{(2)}_{k,k'}(D_n)}_{=0}.$$

Let $\widetilde{\mathcal{K}}_n = (\mathcal{K}_n \setminus \mathcal{K}_{i,n}) \cap \left(\cup_{j\in\partial(\mathcal{K}_n\setminus\mathcal{K}_{i,n})} \mathbb{B}_j(1)\right)$ and note that $\frac{|\widetilde{\mathcal{K}}_n|}{|\mathcal{K}_n|} \to 0$, as $n \to +\infty$. Then,

$$S_3 = \sum_{k\in\mathcal{K}_n\setminus\widetilde{\mathcal{K}}_n} \sum_{k'\in\mathbb{B}_k(1)\cap(\mathcal{K}_n\setminus\mathcal{K}_{i,n})} E^{(2)}_{k,k'}(D_n) + \underbrace{\sum_{k\in\widetilde{\mathcal{K}}_n} \sum_{k'\in\mathbb{B}_k(1)\cap(\mathcal{K}_n\setminus\mathcal{K}_{i,n})} E^{(2)}_{k,k'}(D_n)}_{:=A_3}.$$

Since,

$$\frac{1}{|\mathcal{K}_n|}|A_3| \leq \frac{|\widetilde{\mathcal{K}}_n|}{|\mathcal{K}_n|} \sum_{k \in \mathbb{B}_0(1)} |E^{(2)}_{0,k}(D_n)| \overset{n\to+\infty}{\longrightarrow} 0,$$

we obtain, from Lemma 20,

$$\frac{1}{|\mathcal{K}_n|} S_3 \sim \frac{|\mathcal{K}_n \setminus \widetilde{\mathcal{K}}_n|}{|\mathcal{K}_n|} \sum_{k \in \mathbb{B}_0(1)} E^{(2)}_{0,k}(D_n) \sim \frac{|\mathcal{J}|-1}{|\mathcal{J}|} \sum_{k \in \mathbb{B}_0(1)} E^{(2)}_{0,k}(D).$$

Combining the three terms S_1, S_2 and S_3, we have, as $n \to +\infty$

$$\mathbf{E}\left(\frac{1}{|\mathcal{K}_n|} \sum_{k \in \mathcal{K}_n} \sum_{k' \in \mathcal{K}_n} W_{i,n,\Delta_k(D_n)}(\Phi;\theta^\star)^2\right) \sim \sum_{k \in \mathbb{B}_0(1)} \left(\frac{1}{|\mathcal{J}|} E^{(1)}_{0,k}(D) + \frac{|\mathcal{J}|-1}{|\mathcal{J}|} E^{(2)}_{0,k}(D)\right). \quad (46)$$

When $i \neq j$, there are three main cases in (45), according to $k,k' \in \mathcal{K}_{i,n}$, $k,k' \in \mathcal{K}_{j,n}$, or $k,k' \in \mathcal{K}_n \setminus (\mathcal{K}_{i,n} \cup \mathcal{K}_{j,n})$. As for the case $i=j$ treated before, the other situations involve non-zero correlations on edges sets like $\widetilde{\mathcal{K}}_{i,n}$, which are negligible with respect to $|\mathcal{K}_n|$. The covariance is therefore equivalent, up to $D_n^d|\mathcal{J}|$, to

$$\frac{1}{|\mathcal{K}_n|} \sum_{k,k' \in \mathcal{K}_{i,n}} E^{(12)}_{k,k'}(D_n) + \frac{1}{|\mathcal{K}_n|} \sum_{k,k' \in \mathcal{K}_{j,n}} E^{(12)}_{k,k'}(D_n) + \frac{1}{|\mathcal{K}_n|} \sum_{k,k' \in \mathcal{K}_n \setminus (\mathcal{K}_{i,n} \cup \mathcal{K}_{j,n})} E^{(2)}_{k,k'}(D_n).$$

The simplification occurs as for the case $i=j$ and, since $|\mathcal{K}_{i,n}| = |\mathcal{K}_{j,n}|$, we obtain the asymptotic equivalent for the covariance (45)

$$\frac{1}{D^d|\mathcal{J}|} \sum_{k \in \mathbb{B}_0(1)} \left(\frac{2}{|\mathcal{J}|} E^{(12)}_{0,k}(D) + \frac{|\mathcal{J}|-2}{|\mathcal{J}|} E^{(2)}_{0,k}(D)\right). \quad (47)$$

Finally, from (46) and (47), we deduce that $\underline{\mathbf{\Sigma}}_1(\theta^\star)$, defined in Proposition 4, corresponds to the asymptotic variance of $\left(|\Lambda_{j,n}|^{-1/2} R_{\infty,\Lambda_{j,n}}(\Phi;h,\theta^\star)\right)_{j \in \mathcal{J}}$.

Second step: application of Theorem 21.
We apply Theorem 21 with $\mathbf{Z}_{n,k} = (W_{j,n,\Delta_k(D_n)})_{j \in \mathcal{J}}$, $X_{n,i} = \Phi_{\Delta_i(D_n)}$, $p = |\mathcal{J}|$ and $\underline{\mathbf{\Sigma}} = \underline{\mathbf{\Sigma}}_1(\theta^\star)$, which is a symmetric positive-semidefinite matrix as the limit of a covariance matrix (from the first step of the proof).

The assumption (54) holds from **[N4]**, **[E2(bis)]** and because $D_n \geq D$. Assumptions (i), (ii) and (iii) are direct consequences of **[E2(bis)]**, **[N2]** and Lemma 19. It remains to prove (iv). Assuming $\underline{\mathbf{\Sigma}} = (\Sigma_{ij})$ for $1 \leq i,j \leq p$, from the definition of the Frobenius norm, we have

$$\begin{aligned}\mathbf{E}\left\| |\mathcal{K}_n|^{-1} \sum_{k \in \mathcal{K}_n} \sum_{k' \in \mathbb{B}_k(1) \cap \mathcal{K}_n} \mathbf{Z}_{n,k} {\mathbf{Z}_{n,k'}}^T - \underline{\mathbf{\Sigma}} \right\| \\ \leq \sum_{i=1}^p \sum_{j=1}^p \mathbf{E}\left| |\mathcal{K}_n|^{-1} \sum_{k \in \mathcal{K}_n} \sum_{k' \in \mathbb{B}_k(1) \cap \mathcal{K}_n} W_{i,n,\Delta_k(D_n)} W_{j,n,\Delta_{k'}(D_n)} - \Sigma_{ij} \right|. \quad (48)\end{aligned}$$

Let us first assume that $i \neq j$ are fixed and denote $Y_{n,k}(D_n) = W_{i,n,\Delta_k(D_n)}$, $S^k_n(D_n) = \sum_{k' \in \mathbb{B}_k(1) \cap \mathcal{K}_n} W_{j,n,\Delta_{k'}(D_n)}$. We have

$$\begin{aligned}\mathbf{E}\left| |\mathcal{K}_n|^{-1} \sum_{k \in \mathcal{K}_n} \sum_{k' \in \mathbb{B}_k(1) \cap \mathcal{K}_n} W_{i,n,\Delta_k(D_n)} W_{j,n,\Delta_{k'}(D_n)} - \Sigma_{ij} \right| &= |\mathcal{K}_n|^{-1} \mathbf{E}\left| \sum_{k \in \mathcal{K}_n} Y_{n,k}(D_n) S^k_n(D_n) - \Sigma_{ij} \right| \\ &\leq E_1 + E_2 + E_3 + E_4,\end{aligned}$$

where

$$E_1 = \frac{|\mathcal{K}_{i,n}|}{|\mathcal{K}_n|}\mathbf{E}\left||\mathcal{K}_{i,n}|^{-1}\sum_{k\in\mathcal{K}_{i,n}}(Y_{n,k}(D_n)S_n^k(D_n)-\mathbf{E}(Y_{n,k}(D_n)S_n^k(D_n)))\right|,$$

$$E_2 = \frac{|\mathcal{K}_{j,n}|}{|\mathcal{K}_n|}\mathbf{E}\left||\mathcal{K}_{j,n}|^{-1}\sum_{k\in\mathcal{K}_{j,n}}(Y_{n,k}(D_n)S_n^k(D_n)-\mathbf{E}(Y_{n,k}(D_n)S_n^k(D_n)))\right|,$$

$$\begin{aligned}E_3 &= \frac{|\mathcal{K}_n\setminus(\mathcal{K}_{i,n}\cup\mathcal{K}_{j,n})|}{|\mathcal{K}_n|}\times\\ &\quad\mathbf{E}\left||\mathcal{K}_n\setminus(\mathcal{K}_{i,n}\cup\mathcal{K}_{j,n})|^{-1}\sum_{k\in\mathcal{K}_n\setminus(\mathcal{K}_{i,n}\cup\mathcal{K}_{j,n})}(Y_{n,k}(D_n)S_n^k(D_n)-\mathbf{E}(Y_{n,k}(D_n)S_n^k(D_n)))\right|,\end{aligned}$$

$$\begin{aligned}E_4 = \Bigg|&\frac{|\mathcal{K}_{i,n}|}{|\mathcal{K}_n|}\sum_{k\in\mathcal{K}_{i,n}}\mathbf{E}(Y_{n,k}(D_n)S_n^k(D_n))+\frac{|\mathcal{K}_{i,n}|}{|\mathcal{K}_n|}\sum_{k\in\mathcal{K}_{j,n}}\mathbf{E}(Y_{n,k}(D_n)S_n^k(D_n))\\ &+\frac{|\mathcal{K}_n\setminus(\mathcal{K}_{i,n}\cup\mathcal{K}_{j,n})|}{|\mathcal{K}_n|}\sum_{k\in\mathcal{K}_n\setminus(\mathcal{K}_{i,n}\cup\mathcal{K}_{j,n})}\mathbf{E}(Y_{n,k}(D_n)S_n^k(D_n))-\Sigma_{ij}\Bigg|.\end{aligned}$$

The first three terms E_1, E_2 and E_3 can be handled similarly. Let us focus on E_1:

$$\begin{aligned}\frac{|\mathcal{K}_n|}{|\mathcal{K}_{i,n}|}E_1 \le\ & |\mathcal{K}_{i,n}|^{-1}\sum_{k\in\mathcal{K}_{i,n}}\mathbf{E}\left|Y_{n,k}(D_n)S_n^k(D_n)-Y_{n,k}(D)S_n^k(D)\right|\\ &+|\mathcal{K}_{i,n}|^{-1}\mathbf{E}\left|\sum_{k\in\mathcal{K}_{i,n}}(Y_{n,k}(D)S_n^k(D)-\mathbf{E}(Y_{n,k}(D)S_n^k(D)))\right|\\ &+|\mathcal{K}_{i,n}|^{-1}\sum_{k\in\mathcal{K}_{i,n}}\left|\mathbf{E}\left(Y_{n,k}(D)S_n^k(D)\right)-\mathbf{E}\left(Y_{n,k}(D_n)S_n^k(D_n)\right)\right|.\end{aligned}\tag{49}$$

Up to the edge effects which are negligible with respect to $|\mathcal{K}_{i,n}|$, $\left(Y_{n,k}(D)S_n^k(D)\right)_k$ is stationary when $k\in\mathcal{K}_{i,n}$, since in this case, from (41), $W_{i,n,\Delta_k(D)}=W^{(1)}_{\Delta_k(D)}$ does not depend on n. Therefore the second term in (49) tends to 0 by the mean ergodic theorem. For a fixed n, we have also by stationarity (up to the edge effects)

$$\begin{aligned}\mathbf{E}\left|Y_{n,k}(D_n)S_n^k(D_n)-Y_{n,k}(D)S_n^k(D)\right| &= \mathbf{E}\left|Y_{n,0}(D_n)S_n^0(D_n)-Y_{n,0}(D)S_n^0(D)\right|\\ &= \mathbf{E}\left|W^{(1)}_{\Delta_0(D_n)}W^{(2)}_{\overline{\Delta}_0(D_n)}-W^{(1)}_{\Delta_0(D)}W^{(2)}_{\overline{\Delta}_0(D)}\right|,\end{aligned}$$

where $\overline{\Delta}_0(D_n)=\cup_{k'\in\mathbb{B}_0(1)}\Delta_{k'}(D_n)$. From Lemma 20, this term tends to 0, therefore the first term in (49) asymptotically vanishes. The same argument shows that the third term in (49) also tends to 0 as $n\to\infty$. As a consequence, E_1 tends to 0.

The same decomposition as in (49) may be done for E_2 and E_3, which leads by similar arguments to $E_2\to 0$ and $E_3\to 0$. The last term E_4 involves the difference between Σ_{ij} and its empirical counterpart. The same calculations as in the first step of the proof shows that $E_4\to 0$.

Therefore, we have proved that the terms in the double-sum (48) corresponding to $i\neq j$ asymptotically vanish. The same result can be proved similarly when $i=j$. Thus assumption (iv) in Theorem 21 holds and the convergence in law is deduced.

8.4 Proof of Proposition 6

We can decompose Λ_n in the following way :

$$\Lambda_n := \bigcup_{k \in \mathcal{K}_n} \Delta_k(D_n)$$

where the Δ_k's are disjoint cubes with side-length D_n and $\mathcal{K}_n \subset \mathbb{Z}^d$ satisfies $|\mathcal{K}_n| = |\Lambda_n| D_n^{-d}$. Similarly as in the proof of Proposition 4, we choose

$$D_n = \frac{|\Lambda_n|^{1/d}}{\left\lfloor \frac{|\Lambda_n|^{1/d}}{D} \right\rfloor},$$

which implies $D_n \to D$ when $n \to \infty$ and guarantees $D \leq D_n \leq 2D$ as soon as $|\Lambda_n| \geq D^d$.

From Proposition 3 and under Assumption **[E2(bis)]**, for all $i = 1, \ldots, s$,

$$\begin{aligned} |\Lambda_n|^{-1/2} R_{\Lambda_n}\left(\Phi; h_i, \widehat{\theta}_n(\Phi)\right) &= |\Lambda_n|^{-1/2} R_{\infty,\Lambda_n}(\Phi; h_i, \theta^\star) + o_P(1) \\ &= |\Lambda_n|^{-1/2} \left(I_{\Lambda_n}(\Phi; h_i, \theta^\star) - \mathbf{U}_{\Lambda_n}(\Phi; \theta^\star)^T \mathcal{E}(h_i; \theta^\star) \right) + o_P(1) \\ &= \frac{1}{D_n^{d/2}} \frac{1}{|\mathcal{K}_n|^{1/2}} \sum_{k \in \mathcal{K}_n} W_{\Delta_k(D_n)}(\Phi; h_i, \theta^\star) + o_P(1), \end{aligned}$$

where for any $\varphi \in \Omega$

$$W_{\Delta_k(D_n)}(\varphi; h_i, \theta^\star) := I_{\Delta_k(D_n)}(\varphi; h_i, \theta^\star) + \mathbf{U}_{\Delta_k(D_n)}(\varphi; \theta^\star)^T \mathcal{E}(h_i; \theta^\star).$$

We apply Theorem 21 in the simpler case when $f_{n,k} = f$ for all $n \in \mathbb{N}$ and all $k \in \mathcal{K}_n$. If $D_n = D$ for all n, this framework would reduce to a stationary setting similar to Theorem 2.1 in Jensen and Künsch (1994). But as Λ_n is allowed to increase continuously up to $\mathbb{R}^d$, $D_n \equiv D$ is impossible. We will therefore apply Theorem 21 in Appendix A with $\mathbf{Z}_{n,k} = \left(W_{\Delta_k(D_n)}(\Phi; h_j, \theta^\star)\right)_{j=1\ldots s}$, $X_{n,i} = \Phi_{\Delta_i(D_n)}$ and $p = s$.

Let us first compute the covariance matrix of $\left(|\Lambda_n|^{-1/2} R_{\infty,\Lambda_n}(\Phi; h_i, \theta^\star)\right)_{i=1,\ldots,s}$. By the same calculations as for the term S_1 in the proof of Proposition 4, we obtain

$$\begin{aligned} &cov\left(|\Lambda_n|^{-1/2} R_{\infty,\Lambda_n}(\Phi;\ (h_i, \theta^\star), |\Lambda_n|^{-1/2} R_{\infty,\Lambda_n}(\Phi; h_j, \theta^\star) \right) \\ &\qquad = \frac{1}{D_n^d} \mathbf{E}\left(\frac{1}{|\mathcal{K}_n|} \sum_{k \in \mathcal{K}_n} \sum_{k' \in \mathcal{K}_n} W_{\Delta_k(D_n)}(\Phi; h_i, \theta^\star) W_{\Delta_{k'}(D_n)}(\Phi; h_j, \theta^\star) \right) \\ &\qquad \sim \frac{1}{D^d} \sum_{k \in \mathbb{B}_0(1)} \mathbf{E}\left(W_{\Delta_0(D)}(\Phi; h_i, \theta^\star) W_{\Delta_k(D)}(\Phi; h_j, \theta^\star) \right). \end{aligned} \tag{50}$$

The asymptotic covariance matrix is thus $\underline{\mathbf{\Sigma}}_2(\theta^\star)$ defined in Proposition 6. We can now apply Theorem 21 in the appendix with $\underline{\mathbf{\Sigma}} = \underline{\mathbf{\Sigma}}_2(\theta^\star)$. The assumption (54) hods because $D_n \geq D$ and from **[N4]** and **[E2(bis)]**. The assumptions (i), (ii) and (iii) follow from **[E2(bis)]**, **[N2]** and Lemma 19. Assumption (iv) may be checked easily as in the second step of the proof of Proposition 4, by using (50).

8.5 Proof of Lemma 8

For simplicity, let $\mathbf{Y}_\Lambda := \mathbf{Y}_\Lambda(\Phi; \theta)$. Let us denote $\overline{\Delta}(\delta, D^\vee) := \cup_{|j| \leq \lceil D^\vee/\delta \rceil} \Delta_j(\delta)$. From the additivity property of $\mathbf{Y}$, proving Lemma 8 reduces to prove that for any $\delta > 0$ and any $D^\vee \geq D$, $D^d A(\delta, D^\vee) = \delta^d A(D, D)$ where

$$A(\delta, D^\vee) := \mathbf{E}\left(\mathbf{Y}_{\Delta_0(\delta)} \mathbf{Y}_{\overline{\Delta}(\delta, D^\vee)}{}^T \right).$$

Since $D^\vee \geq D$, we can write $\overline{\Delta}(\delta, D^\vee) = \overline{\Delta}(\delta, D) \cup \Delta'$, where $\Delta' \subset \left(\overline{\Delta}(\delta, D)\right)^c$. From the locality assumption, $\mathbf{Y}_{\Delta'}$ is only a function of $\Phi_{\Delta_0^c(\delta)}$. So

$$\mathbf{E}\left(\mathbf{Y}_{\Delta_0(\delta)}\mathbf{Y}_{\Delta'}{}^T\right) = \mathbf{E}\left(\mathbf{E}\left(\mathbf{Y}_{\Delta_0(\delta)}\mathbf{Y}_{\Delta'}{}^T|\Phi_{\Delta_0^c(\delta)}\right)\right) = \mathbf{E}\left(\mathbf{E}\left(\mathbf{Y}_{\Delta_0(\delta)}|\Phi_{\Delta_0^c(\delta)}\right)\mathbf{Y}_{\Delta'}{}^T\right) = 0, \tag{51}$$

which yields $A(\delta, D^\vee) = A(\delta, D)$. By denoting $A(\delta) := A(\delta, D^\vee) = A(\delta, D)$ and $\overline{\Delta}(\delta) := \overline{\Delta}(\delta, D)$, we must prove $D^d A(\delta) = \delta^d A(D)$.

Let us first assume $\delta = kD$ with $k \in \mathbb{N}$. We may write $\overline{\Delta}(kD) = (\Delta_0(kD) \oplus D) \cup \Delta'$ and may assert that $\mathbf{Y}_{\Delta'}$ depends only on a function of $\Phi_{\Delta_0^c(kD)}$. By a similar argument as in (51), we obtain $A(\delta) = \mathbf{E}(\mathbf{Y}_{\Delta_0(kD)}\mathbf{Y}_{\Delta_0(kD)\oplus D}{}^T)$. From the disjoint decomposition $\Delta_0(kD) = \cup_{j\in\mathcal{K}}\Delta_j(D)$ where $|\mathcal{K}| = k^d$, we have, by the same decorrelation argument as above and by stationarity,

$$\begin{aligned}
A(\delta) &= \sum_{j\in\mathcal{K}} \mathbf{E}(\mathbf{Y}_{\Delta_j(D)}\mathbf{Y}_{\Delta_0(kD)\oplus D}{}^T) = \sum_{j\in\mathcal{K}} \mathbf{E}(\mathbf{Y}_{\Delta_j(D)}\mathbf{Y}_{\Delta_j(D)\oplus D}{}^T) \\
&= k^d \mathbf{E}(\mathbf{Y}_{\Delta_0(D)}\mathbf{Y}_{\Delta_0(D)\oplus D}{}^T) = \frac{\delta^d}{D^d} A(D).
\end{aligned}$$

Let us now assume $D = k\delta$ with $k \in \mathbb{N}$. First notice that in this case $\overline{\Delta}(D) = \overline{\Delta}(\delta) \oplus \frac{D}{2}(1-1/k)$. The following decomposition holds: $\Delta_0(D) = \cup_{j\in\mathcal{K}}\Delta_j(\delta)$ where $|\mathcal{K}| = k^d$. For any $j \in \mathcal{K}$, $|j| \leq \frac{D}{2}(1-1/k)$, so $\overline{\Delta}(D)$ contains any translation of the set $\overline{\Delta}(\delta)$ with respect to j. Let us denote this translated set by $\tau_j\overline{\Delta}(\delta)$. From the same decorrelation argument as above and by stationarity, we have

$$A(D) = \sum_{j\in\mathcal{K}} \mathbf{E}(\mathbf{Y}_{\Delta_j(\delta)}\mathbf{Y}_{\overline{\Delta}(D)}{}^T) = \sum_{j\in\mathcal{K}} \mathbf{E}(\mathbf{Y}_{\Delta_j(\delta)}\mathbf{Y}_{\tau_j\overline{\Delta}(\delta)}{}^T) = \sum_{j\in\mathcal{K}} \mathbf{E}(\mathbf{Y}_{\Delta_0(\delta)}\mathbf{Y}_{\overline{\Delta}(\delta)}{}^T) = \frac{D^d}{\delta^d} A(\delta). \tag{52}$$

Let us now consider the case $D/\delta = k'/k$, where $(k, k') \in \mathbb{N}^2$. Let $\delta' = \delta/k$, then $D = k'\delta'$ and according to (52), ${\delta'}^d A(D) = D^d A(\delta')$. In the same way as we have proved $D^d A(\delta) = \delta^d A(D)$ when $\delta = kD$, this is not difficult to show that for any $\delta = k\delta'$ with $\delta' \leq D$, ${\delta'}^d A(\delta) = \delta^d A(\delta')$. As a consequence when $D/\delta = k'/k$, we obtain

$$A(D) = \frac{D^d}{{\delta'}^d} A(\delta') = \frac{D^d}{\delta^d} A(\delta). \tag{53}$$

In the general case, one may find a sequence of rational numbers $(q_n)_{n\in\mathbb{N}}$ which converges to D/δ. Let $\delta_n = q_n D$, we have from (53), $A(D) = \frac{D^d}{\delta_n^d} A(\delta_n)$. Since we have assumed $\mathbf{E}(\mathbf{Y}_{\Gamma_n}^2) \to 0$ when $\Gamma_n \to 0$, the additivity of $\mathbf{Y}$ and $\delta_n \to \delta$ yield

$$A(\delta_n) = \mathbf{E}\left(\mathbf{Y}_{\Delta_0(\delta_n)}\mathbf{Y}_{\overline{\Delta}(\delta_n)}{}^T\right) \to \mathbf{E}\left(\mathbf{Y}_{\Delta_0(\delta)}\mathbf{Y}_{\overline{\Delta}(\delta)}{}^T\right) = A(\delta)$$

as n goes to infinity. Therefore, the identity (53) holds for any $\delta > 0$, which concludes the proof.

8.6 Proof of Proposition 7

The proof follows arguments presented by Jensen and Künsch in Jensen and Künsch (1994). Let $C_n(\overline{\delta}) = [-n\overline{\delta} - \overline{\delta}/2, n\overline{\delta} + \overline{\delta}/2]^d$, so $C_n(\overline{\delta}) = \cup_{k\in\mathcal{K}_n}\Delta_k(\overline{\delta})$, where $\mathcal{K}_n = [-n, n]^d \cap \mathbb{Z}^d$ and $\Delta_k(\overline{\delta})$ is the cube centered at $k\overline{\delta}$ with side-length $\overline{\delta}$. We have

$$\begin{aligned}
Var\left(|C_n(\overline{\delta})|^{-1/2}\mathbf{Y}_{C_n(\overline{\delta})}(\Phi;\theta^\star)\right) &= |C_n(\overline{\delta})|^{-1} \sum_{i,j\in\mathcal{K}_n} \mathbf{E}\left(\mathbf{Y}_{\Delta_i(\overline{\delta})}(\Phi;\theta^\star)\,\mathbf{Y}_{\Delta_j(\overline{\delta})}(\Phi;\theta^\star)^T\right) \\
&= |C_n(\overline{\delta})|^{-1} \sum_{i\in\mathcal{K}_n} \sum_{j\in\mathbb{B}_i(\lceil \frac{D}{\delta} \rceil)\cap\mathcal{K}_n} \mathbf{E}\left(\mathbf{Y}_{\Delta_i(\overline{\delta})}(\Phi;\theta^\star)\,\mathbf{Y}_{\Delta_j(\overline{\delta})}(\Phi;\theta^\star)^T\right).
\end{aligned}$$

Since $|C_n(\overline{\delta})| = \overline{\delta}^d |\mathcal{K}_n|$, from the ergodic theorem,

$$Var\left(|C_n(\overline{\delta})|^{-1/2}\mathbf{Y}_{C_n(\overline{\delta})}(\Phi;\theta^\star)\right) \longrightarrow \overline{\delta}^{-d} \sum_{|k|\leq \lceil \frac{D}{\overline{\delta}} \rceil} \mathbf{E}\left(\mathbf{Y}_{\Delta_0(\overline{\delta})}(\Phi;\theta^\star)\,\mathbf{Y}_{\Delta_k(\overline{\delta})}(\Phi;\theta^\star)^T\right)$$

which is $\underline{\mathbf{M}}(\theta^\star)$ by Lemma 8.

Therefore, to prove that $\underline{\mathbf{M}}(\theta^\star)$ is positive-definite, it is sufficient to prove that the covariance matrix $Var\left(|C_n(\overline{\delta})|^{-1/2}\mathbf{Y}_{C_n(\overline{\delta})}(\Phi;\theta^\star)\right)$ is positive-definite for n large enough. Let $\mathbf{x} \in \mathbb{R}^q \setminus \{0\}$, we must show that

$$V := \mathbf{x}^T Var\left(|C_n(\overline{\delta})|^{-1/2}\mathbf{Y}_{C_n(\overline{\delta})}(\Phi;\theta^\star)\right)\mathbf{x} > 0.$$

Since, for two random variables X, X' with finite variance

$$Var(X) = \mathbf{E}(Var(X|X')) \;+\; Var(\mathbf{E}(X|X')) \geq \mathbf{E}(Var(X|X')),$$

we have, by denoting $L := \left(2\left\lceil \frac{D}{\overline{\delta}}\right\rceil + 1\right)\mathbb{Z}^d$,

$$\begin{aligned}
V &\geq |C_n(\overline{\delta})|^{-1}\,\mathbf{E}\left(Var\left(\mathbf{x}^T\mathbf{Y}_{C_n(\overline{\delta})}(\Phi;\theta^\star)|\;\Phi_{\Delta_k(\overline{\delta})}, k \notin L\right)\right)\\
&= |C_n(\overline{\delta})|^{-1}\mathbf{x}^T\mathbf{E}\Bigg(Var\Bigg(\sum_{\ell\in L\cap\mathcal{K}_n} \underbrace{\sum_{i\in\mathbb{B}_\ell(\lceil\frac{D}{\overline{\delta}}\rceil)\cap\mathcal{K}_n} \mathbf{Y}_{\Delta_i(\overline{\delta})}(\Phi;\theta^\star)}_{:=\mathbf{S}_{\ell,n}(\Phi)} \,|\; \Phi_{\Delta_k(\overline{\delta})}, k \notin L\Bigg)\Bigg)\mathbf{x}
\end{aligned}$$

Note that from the locality property, $\mathbf{S}_{\ell,n}(\Phi)$ depends only on $\Phi_{\Delta_j(\overline{\delta})}$ for $j \in \mathbb{B}_\ell\left(2\left\lceil\frac{D}{\overline{\delta}}\right\rceil\right)$. Therefore, conditionally on $\Phi_{\Delta_k(\overline{\delta})}, k \notin L$, the variables $\mathbf{S}_{\ell,n}(\Phi)$ and $\mathbf{S}_{\ell',n}(\Phi)$ (for $\ell \neq \ell'$) are independent. Now, let $\overline{\Delta}(\overline{\delta}) := \cup_{|i|\leq\lceil\frac{D}{\overline{\delta}}\rceil}\Delta_i(\overline{\delta})$, from the stationarity we have for n large enough

$$\begin{aligned}
V &\geq |C_n(\overline{\delta})|^{-1}\mathbf{x}^T\sum_{\ell\in L\cap\mathcal{K}_n}\mathbf{E}\Big(Var\Big(\mathbf{S}_{\ell,n}(\Phi)|\;\Phi_{\Delta_k(\overline{\delta})}, k\notin L\Big)\Big)\mathbf{x}\\
&\geq \frac{\overline{\delta}^{-d}}{2}\frac{|L\cap\mathcal{K}_n|}{|\mathcal{K}_n|}\times\mathbf{E}\left(Var\left(\mathbf{x}^T\mathbf{Y}_{\overline{\Delta}(\overline{\delta})}(\Phi;\theta^\star)\,|\Phi_{\Delta_k(\overline{\delta})}, 1\leq|k|\leq 2\left\lceil\frac{D}{\overline{\delta}}\right\rceil\right)\right)\\
&\geq \kappa(\overline{\delta}, D, d)\times\mathbf{E}\left(Var\left(\mathbf{x}^T\mathbf{Y}_{\overline{\Delta}(\overline{\delta})}(\Phi;\theta^\star)\,|\Phi_{\Delta_k(\overline{\delta})}, 1\leq|k|\leq 2\left\lceil\frac{D}{\overline{\delta}}\right\rceil\right)\right),
\end{aligned}$$

where $\kappa(\overline{\delta}, D, d)$ is a positive constant. Assume there exists some positive constant c such that $P_{\theta^\star}-$a.s. $\mathbf{x}^T\mathbf{Y}_{\overline{\Delta}(\overline{\delta})}(\Phi;\theta^\star) = c$ when the variables $\Phi_{\Delta_k(\overline{\delta})}, 1 \leq |k| \leq 2\left\lceil\frac{D}{\overline{\delta}}\right\rceil$ are fixed to belong to B, where $B \in \mathcal{F}$ is involved in **[PD]**. It follows that for any $\varphi_i \in A_i$ for $i = 0,\ldots,\ell$ (with $\ell \geq 1$), where the A_i's come from **[PD]**, $\mathbf{x}^T\left(\mathbf{Y}_{\overline{\Delta}(\overline{\delta})}(\varphi_i;\theta^\star) - \mathbf{Y}_{\overline{\Delta}(\overline{\delta})}(\varphi_0;\theta^\star)\right) = 0$. Since for all $(\varphi_0,\ldots,\varphi_\ell) \in A_0\times\ldots\times A_\ell$, the matrix with entries $\left(\mathbf{Y}_{\overline{\Delta}(\overline{\delta})}(\varphi_i;\theta^\star)\right)_j - \left(\mathbf{Y}_{\overline{\Delta}(\overline{\delta})}(\varphi_0;\theta^\star)\right)_j$ is assumed to be injective, this leads to $\mathbf{x} = 0$ and hence to some contradiction. Therefore, when the variables $\Phi_{\Delta_k(\overline{\delta})}$, $1 \leq |k| \leq 2\left\lceil\frac{D}{\overline{\delta}}\right\rceil$ are for example assumed to belong to B, the variable $\mathbf{x}^T\mathbf{Y}_{\overline{\Delta}(\overline{\delta})}(\Phi;\theta^\star)$ is almost surely not a constant and so $V > 0$, which proves that $\underline{\mathbf{M}}(\theta^\star)$ is a symmetric positive-definite matrix.

8.7 Proof of Proposition 9

Since for any $\varphi \in \Omega$, $\widehat{\underline{\mathbf{M}}}_n(\varphi;\cdot,\delta,D^\vee)$ is continuous in a neighborhood $\mathcal{V}(\theta^\star)$ of $\theta^\star$ and according to **[E1]**, it is sufficient to prove that for any $\theta \in \mathcal{V}(\theta^\star)$, $\widehat{\underline{\mathbf{M}}}_n(\Phi;\theta,\delta_n,D^\vee)$ converges in probability towards $\underline{\mathbf{M}}(\theta)$.

We choose the sequence δ_n as follows :

$$\delta_n = \frac{|\Lambda_n|^{1/d}}{\left\lfloor \frac{|\Lambda_n|^{1/d}}{\delta} \right\rfloor},$$

which guarantees $\delta_{n_0} = \delta$, since $|\Lambda_{n_0}|\delta^{-d} \in \mathbb{N}$, $\delta \le \delta_n \le 2\delta$ for n sufficiently large, and $\delta_n \to \delta$ as $n \to \infty$. This choice allows us to consider, for all $n \in \mathbb{N}$, the decomposition $\Lambda_n = \cup_{k\in\mathcal{K}_n}\Delta_k(\delta_n)$, where the $\Delta_k(\delta_n)$'s are disjoint cubes with side-length δ_n and centered at $k\delta_n$. Moreover, since $\delta_n \ge \delta$ and $\delta_n \to \delta$, we have $\left\lceil \frac{D^\vee}{\delta_n} \right\rceil = \left\lceil \frac{D^\vee}{\delta} \right\rceil$ when n is large enough, which is assumed in the sequel of the proof.

Let $\widetilde{\mathcal{K}}_n := \mathcal{K}_n \cap \left(\cup_{j\in\partial\mathcal{K}_n} \mathbb{B}_j\left(\left\lceil \frac{D^\vee}{\delta} \right\rceil\right)\right)$. Since $|\Lambda_n| = \delta_n^d|\mathcal{K}_n|$, we have

$$\left|\delta_n^d\widehat{\underline{\mathbf{M}}}_n(\Phi;\theta,\delta_n,D^\vee) - \delta^d\underline{\mathbf{M}}(\theta)\right| \le X_1 + X_2 + X_3 + X_4,$$

where by setting $\overline{\Delta}_k(\tau) = \cup_{j\in\mathbb{B}_k(\lceil \frac{D^\vee}{\delta}\rceil)}\Delta_j(\tau)$ (for some $\tau > 0$),

$$\begin{aligned}
X_1 &= \left| |\mathcal{K}_n|^{-1} \sum_{k\in\widetilde{\mathcal{K}}_n} \sum_{j\in\mathbb{B}_k(\lceil \frac{D^\vee}{\delta}\rceil)\cap\mathcal{K}_n} \widehat{\mathbf{Y}}_{n,\Delta_k(\delta_n)}(\Phi;\theta)\,\widehat{\mathbf{Y}}_{n,\Delta_j(\delta_n)}(\Phi;\theta)^T \right| \\
X_2 &= \left| |\mathcal{K}_n|^{-1} \sum_{k\in\mathcal{K}_n\setminus\widetilde{\mathcal{K}}_n} \left(\widehat{\mathbf{Y}}_{n,\Delta_k(\delta_n)}(\Phi;\theta)\,\widehat{\mathbf{Y}}_{n,\overline{\Delta}_k(\delta_n)}(\Phi;\theta)^T - \mathbf{Y}_{\Delta_k(\delta_n)}(\Phi;\theta)\,\mathbf{Y}_{\overline{\Delta}_k(\delta_n)}(\Phi;\theta)^T \right) \right| \\
X_3 &= \left| |\mathcal{K}_n|^{-1} \sum_{k\in\mathcal{K}_n\setminus\widetilde{\mathcal{K}}_n} \left(\mathbf{Y}_{n,\Delta_k(\delta_n)}(\Phi;\theta)\,\mathbf{Y}_{n,\overline{\Delta}_k(\delta_n)}(\Phi;\theta)^T - \mathbf{Y}_{\Delta_k(\delta)}(\Phi;\theta)\,\mathbf{Y}_{\overline{\Delta}_k(\delta)}(\Phi;\theta)^T \right) \right| \\
X_4 &= \left| |\mathcal{K}_n|^{-1} \sum_{k\in\mathcal{K}_n\setminus\widetilde{\mathcal{K}}_n} \mathbf{Y}_{\Delta_k(\delta)}(\Phi;\theta)\,\mathbf{Y}_{\overline{\Delta}_k(\delta)}(\Phi;\theta)^T - \mathbf{E}\left(\mathbf{Y}_{\Delta_0(\delta)}(\Phi;\theta)\,\mathbf{Y}_{\overline{\Delta}_0(\delta)}(\Phi;\theta)^T \right) \right|.
\end{aligned}$$

We have from the additivity and the stationartiy of $\widehat{\mathbf{Y}}_n$,

$$\begin{aligned}
\mathbf{E}|X_1| &\le |\mathcal{K}_n|^{-1} \sum_{k\in\widetilde{\mathcal{K}}_n} \sum_{j\in\mathbb{B}_k(\lceil \frac{D^\vee}{\delta}\rceil)\cap\mathcal{K}_n} \mathbf{E}\left| \widehat{\mathbf{Y}}_{n,\Delta_k(\delta_n)}(\varphi;\theta)\,\widehat{\mathbf{Y}}_{n,\Delta_j(\delta_n)}(\varphi;\theta)^T \right| \\
&\le \frac{|\widetilde{\mathcal{K}}_n|}{|\mathcal{K}_n|}\mathbf{E}\left| \widehat{\mathbf{Y}}_{n,\Delta_0(\delta_n)}(\Phi;\theta)\widehat{\mathbf{Y}}_{n,\overline{\Delta}_0(\delta_n)}(\Phi;\theta)^T \right|,
\end{aligned}$$

which tends to 0 as $n \to +\infty$, because $\frac{|\widetilde{\mathcal{K}}_n|}{|\mathcal{K}_n|} \to 0$ and $\delta \le \delta_n \le 2\delta$. Therefore, X_1 converges in probability to 0. The second term converges also to 0 in probability from the additivity of $\mathbf{Y}$ and $\widehat{\mathbf{Y}}$ and from (16). The expectation of the third term converges to 0 by following the proof of Lemma 20. Finally, from the stationarity of $\mathbf{Y}$ and since $|\mathcal{K}_n| \sim |\mathcal{K}_n \setminus \widetilde{\mathcal{K}}_n|$, the mean ergodic theorem applies to $\mathbf{E}|X_4|$, which, in particular, shows that $X_4 \to 0$ in probability. This proves that

$$\delta_n^d\widehat{\underline{\mathbf{M}}}_n(\Phi;\theta,\delta_n,D^\vee) \longrightarrow \delta^d\underline{\mathbf{M}}(\theta),$$

in probability, as $n \to \infty$. Since δ_n is a deterministic sequence converging to δ, the conclusion of Proposition 9 follows.

A Central Limit Theorem

The following result is a central limit theorem for conditionnally centered random fields. It generalizes Theorem 2.1 in Jensen and Künsch (1994) to a non-stationary and non-ergodic setting. A general result has been proved by Comets and Janzura (1998) for self normalized sums, provided a fourth moment condition. Our result is in the same spirit but it is proved for triangular array and without self-normalization, which is well-adapted to the residuals process framework. This allows in particular to avoid the fourth moment assumption.

Theorem 21. *Let $X_{n,i}$, $n \in \mathbb{N}$, $i \in \mathbb{Z}^d$, be a triangular array field in a measurable space S. For $n \in \mathbb{N}$, let $\mathcal{K}_n \subset \mathbb{Z}^d$ and for $k \in \mathcal{K}_n$, assume*

$$\mathbf{Z}_{n,k} = f_{n,k}\left(X_{n,k+i},\ i \in \mathcal{I}_0\right), \tag{54}$$

where $\mathcal{I}_0 = \{i \in \mathbb{Z}^d,\ |i| \leq 1\}$ and $f_{n,k} : S^{\mathcal{I}_0} \to \mathbb{R}^p$. Let $\mathbf{S}_n = \sum_{k\in\mathcal{K}_n} \mathbf{Z}_{n,k}$. If

(i) $c_3 := \sup_{n\in\mathbb{N}} \sup_{k\in\mathcal{K}_n} \mathbf{E}|\mathbf{Z}_{n,k}|^3 < \infty$,

(ii) $\forall n \in \mathbb{N}$, $\forall k \in \mathcal{K}_n$, $\mathbf{E}(\mathbf{Z}_{n,k}|X_{n,j},\ j \neq k) = 0$,

(iii) $|\mathcal{K}_n| \to +\infty$ *as* $n \to \infty$,

(iv) There exists a symmetric matrix $\underline{\mathbf{\Sigma}} \geq 0$ *such that*

$$\mathbf{E}\left\|\,|\mathcal{K}_n|^{-1} \sum_{k\in\mathcal{K}_n} \sum_{j\in\mathbb{B}_k(1)\cap\mathcal{K}_n} \mathbf{Z}_{n,k}\mathbf{Z}_{n,j}{}^T - \underline{\mathbf{\Sigma}}\right\| \to 0,$$

then $|\mathcal{K}_n|^{-1/2}\mathbf{S}_n \xrightarrow{d} \mathcal{N}\left(0, \underline{\mathbf{\Sigma}}\right)$ *as* $n \to \infty$.

Proof.

Let us first assume that $\underline{\mathbf{\Sigma}}$ is a positive-definite matrix (i.e. $\underline{\mathbf{\Sigma}} > 0$). According to the Stein's method (see also Bolthausen (1982)), it suffices to prove that, for all $\mathbf{e} \in \mathbb{R}^p$ such that $\|\mathbf{e}\| = 1$ and for all $\lambda \in \mathbb{R}$,

$$\mathbf{E}\left(\left(i\lambda - \mathbf{e}^T|\mathcal{K}_n|^{-1/2}\underline{\mathbf{\Sigma}}^{-1/2}\mathbf{S}_n\right)\exp\left(i\lambda\mathbf{e}^T|\mathcal{K}_n|^{-1/2}\underline{\mathbf{\Sigma}}^{-1/2}\mathbf{S}_n\right)\right) \to 0.$$

Denoting $\mathbf{u} = \lambda\mathbf{e}$, this is equivalent to prove that for all $\mathbf{u} \in \mathbb{R}^p$,

$$\mathbf{E}\left(\underbrace{\left(i\mathbf{u} - |\mathcal{K}_n|^{-1/2}\underline{\mathbf{\Sigma}}^{-1/2}\mathbf{S}_n\right)\exp(i\mathbf{u}^T|\mathcal{K}_n|^{-1/2}\underline{\mathbf{\Sigma}}^{-1/2}\mathbf{S}_n)}_{:=\mathbf{A}}\right) \to 0.$$

We decompose the term $\mathbf{A}$ in the same spirit as in Bolthausen (1982), Jensen and Künsch (1994) and Comets and Janzura (1998). Let us denote by $\underline{\mathbf{I}}_p$ the identity matrix of size p and $\mathbf{S}_n^k = \sum_{j\in\mathbb{B}_k(1)\cap\mathcal{K}_n} \mathbf{Z}_{n,j}$. Noting that $\mathbf{u}^T\underline{\mathbf{\Sigma}}^{-1/2}\mathbf{S}_n^k = \mathbf{S}_n^{k\,T}\underline{\mathbf{\Sigma}}^{-1/2\,T}\mathbf{u}$, the decomposition is $\mathbf{E}(\mathbf{A}) = \mathbf{E}(\mathbf{A}_1 - \mathbf{A}_2 - \mathbf{A}_3)$ where

$$\begin{aligned}
\mathbf{A}_1 &= i\exp(i\mathbf{u}^T|\mathcal{K}_n|^{-1/2}\underline{\mathbf{\Sigma}}^{-1/2}\mathbf{S}_n)\left[\underline{\mathbf{I}}_p - |\mathcal{K}_n|^{-1}\underline{\mathbf{\Sigma}}^{-1/2}\sum_{k\in\mathcal{K}_n}\mathbf{Z}_{n,k}\mathbf{S}_n^{k\,T}\right]\mathbf{u},\\
\mathbf{A}_2 &= \exp(i\mathbf{u}^T|\mathcal{K}_n|^{-1/2}\underline{\mathbf{\Sigma}}^{-1/2}\mathbf{S}_n)|\mathcal{K}_n|^{-1/2}\underline{\mathbf{\Sigma}}^{-1/2}\\
&\qquad\times\sum_{k\in\mathcal{K}_n}\mathbf{Z}_{n,k}\left(1 - \exp(-i\mathbf{u}^T|\mathcal{K}_n|^{-1/2}\underline{\mathbf{\Sigma}}^{-1/2}\mathbf{S}_n^k) - i\mathbf{u}^T|\mathcal{K}_n|^{-1/2}\underline{\mathbf{\Sigma}}^{-1/2}\mathbf{S}_n^k\right),\\
\mathbf{A}_3 &= |\mathcal{K}_n|^{-1/2}\underline{\mathbf{\Sigma}}^{-1/2}\sum_{k\in\mathcal{K}_n}\mathbf{Z}_{n,k}\exp\left[i\mathbf{u}^T|\mathcal{K}_n|^{-1/2}\underline{\mathbf{\Sigma}}^{-1/2}(\mathbf{S}_n - \mathbf{S}_n^k)\right].
\end{aligned}$$

The two last terms $\mathbf{A}_2$ and $\mathbf{A}_3$ can be handled as in Jensen and Künsch (1994): $\mathbf{E}(\mathbf{A}_3) = 0$ by (ii) and $|\mathbf{E}(\mathbf{A}_2)| \to 0$ from the same inequalities therein and the sub-multiplicative property of the Frobenius norm. These inequalities rely on two facts: $\forall x \in \mathbb{R}, |1 - e^{-ix} - ix| \leq x^2/2$ and for all n and for all $(k_1, k_2, k_3) \in \mathcal{K}_n$, $\mathbf{E}(|\mathbf{Z}_{n,k_1}||\mathbf{Z}_{n,k_2}||\mathbf{Z}_{n,k_3}|) \leq \left(\mathbf{E}(|\mathbf{Z}_{n,k_1}|^3)\mathbf{E}(|\mathbf{Z}_{n,k_2}|^3)\mathbf{E}(|\mathbf{Z}_{n,k_3}|^3)\right)^{1/3}$ which is less than c_3 by (i).

For $\mathbf{A}_1$, we cannot use a mean ergodic theorem as in Jensen and Künsch (1994), but Assumption (iv) is sufficient. Indeed,

$$\begin{aligned}\|\mathbf{E}(\mathbf{A}_1)\| &\leq \|\mathbf{u}\|\mathbf{E}\left\|\underline{\mathbf{I}}_p - |\mathcal{K}_n|^{-1}\underline{\mathbf{\Sigma}}^{-1/2} \sum_{k\in\mathcal{K}_n}\sum_{j\in\mathbb{B}_k(1)\cap\mathcal{K}_n} \mathbf{Z}_{n,k}{\mathbf{Z}_{n,j}}^T{\underline{\mathbf{\Sigma}}^{-1/2}}^T\right\| \\ &\leq \|\mathbf{u}\| \left\|\underline{\mathbf{\Sigma}}^{-1/2}\right\|^2 \mathbf{E}\left\||\mathcal{K}_n|^{-1}\sum_{k\in\mathcal{K}_n}\sum_{j\in\mathbb{B}_k(1)\cap\mathcal{K}_n} \mathbf{Z}_{n,k}{\mathbf{Z}_{n,j}}^T - \underline{\mathbf{\Sigma}}\right\|\end{aligned}$$

which tends to 0 by (iv).

Now, if $\underline{\mathbf{\Sigma}}$ is not a positive-definite matrix, one can find an orthonormal basis $(\mathbf{f}_1, \ldots, \mathbf{f}_p)$ of $\mathbb{R}^p$, where the $\mathbf{f}_i$'s are eigenvectors of $\underline{\mathbf{\Sigma}}$. We agree that, if $r < p$ denotes the rank of $\underline{\mathbf{\Sigma}}$, then $(\mathbf{f}_1, \ldots, \mathbf{f}_r)$ is a basis of the image of $\underline{\mathbf{\Sigma}}$, while $(\mathbf{f}_{r+1}, \ldots, \mathbf{f}_p)$ is a basis of its kernel.

Let us denote by $\underline{\mathbf{V}}_{Im}$ the matrix whose columns are $(\mathbf{f}_1, \ldots, \mathbf{f}_r)$ and $\underline{\mathbf{V}}_{Ker}$ the matrix whose columns are $(\mathbf{f}_{r+1}, \ldots, \mathbf{f}_p)$. Similarly, for any $\mathbf{u} \in \mathbb{R}^p$, let us denote by u_i its i-th coordinate in the basis $(\mathbf{f}_1, \ldots, \mathbf{f}_p)$ and $\mathbf{u}_{Im} = (u_1, \ldots, u_r)$, $\mathbf{u}_{Ker} = (u_{r+1}, \ldots, u_p)$. Hence $\mathbf{u} = \underline{\mathbf{V}}_{Im}\mathbf{u}_{Im} + \underline{\mathbf{V}}_{Ker}\mathbf{u}_{Ker}$.

The convergence in law of $|\mathcal{K}_n|^{-1/2}\mathbf{S}_n$ to a Gaussian vector reduces to the convergence of $\mathbf{u}^T|\mathcal{K}_n|^{-1/2}\mathbf{S}_n$ for all $\mathbf{u} \in \mathbb{R}^p$. We have

$$\mathbf{u}^T|\mathcal{K}_n|^{-1/2}\mathbf{S}_n = {\mathbf{u}_{Im}}^T{\underline{\mathbf{V}}_{Im}}^T|\mathcal{K}_n|^{-1/2}\mathbf{S}_n + {\mathbf{u}_{Ker}}^T{\underline{\mathbf{V}}_{Ker}}^T|\mathcal{K}_n|^{-1/2}\mathbf{S}_n. \tag{55}$$

From (iv) and since ${\underline{\mathbf{V}}_{Ker}}^T\underline{\mathbf{\Sigma}}\,\underline{\mathbf{V}}_{Ker} = 0$, we deduce that

$$\mathbf{E}\left\||\mathcal{K}_n|^{-1}\sum_{k\in\mathcal{K}_n}\sum_{j\in\mathbb{B}_k(1)\cap\mathcal{K}_n} {\underline{\mathbf{V}}_{Ker}}^T\mathbf{Z}_{n,k}{\mathbf{Z}_{n,j}}^T\underline{\mathbf{V}}_{Ker}\right\| \longrightarrow 0,$$

which means that ${\underline{\mathbf{V}}_{Ker}}^T|\mathcal{K}_n|^{-1/2}\mathbf{S}_n$ tends to 0 in quadratic mean.

On the other hand, the assumptions of Theorem 21 imply that $(i) - (iv)$ remain true when one replaces $\mathbf{Z}_{n,k}$ by ${\underline{\mathbf{V}}_{Im}}^T\mathbf{Z}_{n,k}$ and $\underline{\mathbf{\Sigma}}$ by ${\underline{\mathbf{V}}_{Im}}^T\underline{\mathbf{\Sigma}}\,\underline{\mathbf{V}}_{Im}$. Since ${\underline{\mathbf{V}}_{Im}}^T\underline{\mathbf{\Sigma}}\,\underline{\mathbf{V}}_{Im}$ is positive-definite, the convergence in law of ${\underline{\mathbf{V}}_{Im}}^T|\mathcal{K}_n|^{-1/2}\mathbf{S}_n$ holds for the same reasons as in the first part of the proof.

Therefore, we have proved that for all $\mathbf{u} \in \mathbb{R}^p$, $\mathbf{u}^T|\mathcal{K}_n|^{-1/2}\mathbf{S}_n \xrightarrow{d} \mathcal{N}\left(0, {\mathbf{u}_{Im}}^T{\underline{\mathbf{V}}_{Im}}^T\underline{\mathbf{\Sigma}}\,\underline{\mathbf{V}}_{Im}\mathbf{u}_{Im}\right)$. It is easy to check that ${\mathbf{u}_{Im}}^T{\underline{\mathbf{V}}_{Im}}^T\underline{\mathbf{\Sigma}}\,\underline{\mathbf{V}}_{Im}\mathbf{u}_{Im} = \mathbf{u}^T\underline{\mathbf{\Sigma}}\mathbf{u}$, which concludes the proof. ■

B Assumption [PD] on two examples

In this section, we focus on the two following models, belonging to the exponential family :

1. **Two-type marked Strauss point process** : $\mathbb{M} = \{1, 2\}$ and for $\theta = \left(\theta_1^1, \theta_1^2, \theta_2^{1,1}, \theta_2^{1,2}, \theta_2^{2,2}\right)^T$, for any $\Lambda \in \mathcal{B}(\mathbb{R}^d)$,

$$V_\Lambda(\varphi;\theta) = \theta_1^1 \underbrace{|\varphi_\Lambda^1|}_{:=v_{\Lambda,1}^1(\varphi)} + \theta_1^2 \underbrace{|\varphi_\Lambda^2|}_{:=v_{\Lambda,1}^2(\varphi)} + \sum_{\substack{m_1,m_2=1\\ m_1\leq m_2}}^{2} \theta_2^{m_1,m_2} \underbrace{\sum_{\substack{\{x_1^{m_1},x_2^{m_2}\}\in\mathcal{P}_2(\varphi)\\ \{x_1^{m_1},x_2^{m_2}\}\cap\Lambda\neq\emptyset}} \mathbf{1}_{[0,D^{m_1,m_2}]}(\|x_2 - x_1\|)}_{:=v_{\Lambda,2}^{m_1,m_2}(\varphi)}.$$

Alternatively,

$$V\left(x^m|\varphi;\theta\right) = \theta_1^m + \sum_{m'=1}^{2} \theta_2^{m,m'} \sum_{y^{m'}\in\varphi} \mathbf{1}_{[D_0^{m_1,m_2},D^{m_1,m_2}]}(\|y-x\|).$$

This process is well-defined when $\theta_2^{m_1,m_2} \geq 0$ and $D_0^{m_1,m_2} = 0$ (inhibition assumption), or when $\theta_2^{m_1,m_2} \in \mathbb{R}^2$ and $D_0^{m_1,m_2} = \delta > 0$ (hard-core assumption), see Proposition 13 for instance. The range of the local energy function equals $D = \max\left(D^{1,1}, D^{1,2}, D^{2,2}\right)$.

2. **Area interaction point process** : $\mathbb{M} = \{0\}$ and for $R > 0$, $\theta = (\theta_1, \theta_2)$ and any $\Lambda \in \mathcal{B}(\mathbb{R}^d)$,

$$V_\Lambda\left(\varphi;\theta\right) = \theta_1|\varphi_\Lambda| + \theta_2 v_2(\varphi_\Lambda), \quad \text{with} \;\; v_2(\varphi_\Lambda) := \left|\cup_{x\in\varphi_\Lambda} B(x,R)\right|.$$

Alternatively,

$$v_1(0|\varphi) := 1, \qquad v_2(0|\varphi) := \left|\cup_{x\in(\varphi_{\mathcal{B}(0,2R)}\cup\{0\})}\mathcal{B}(x,R) \setminus \cup_{x\in\varphi_{\mathcal{B}(0,2R)}}\mathcal{B}(x,R)\right|.$$

This model is well-defined for $\theta \in \mathbb{R}^2$ (see Proposition 13 for instance) and the range of the local energy equals $D = 2R$.

Both these models satisfy **[C]** and **[N1-4]**. The aim of the sequel is to prove Proposition 18, which claims that $\underline{\boldsymbol{\Sigma}}_1(\theta^\star)$ and $\underline{\boldsymbol{\Sigma}}_2(\theta^\star)$, involved respectively in Proposition 4 and 6, are positive-definite for these models, when considering the maximum pseudolikelihood estimate for $\widehat{\theta}_n$ and the two following frameworks

- Framework 1 (for $\underline{\boldsymbol{\Sigma}}_1(\theta^\star)$): we consider the inverse residuals ($h = e^V$). Let us recall that Proposition 17 asserts that **[PD]** fails for the raw residuals ($h = 1$) for both the area-interaction and 2-type marked Strauss models.

- Framework 2 (for $\underline{\boldsymbol{\Sigma}}_2(\theta^\star)$): we consider the family of test functions given for $j = 1, \ldots, s$ and $0 < r_1 < \ldots < r_s < +\infty$ by

$$h_j(x^m, \varphi;\theta) = \mathbf{1}_{[0,r_j]}(d(x^m,\varphi))e^{V(x^m|\varphi;\theta)},$$

  related to parametric and nonparametric estimations of the empty space function at distance r_j.

When considering the MPLE, $R_{\infty,\Lambda}(\varphi;h,\theta^\star)$ is given by (26) with $\mathbf{LPL}^{(1)}$, $\underline{\mathbf{H}}$ and $\mathcal{E}$ respectively given by (24), (25) and (7).

B.1 2-type marked Strauss point process

We only deal with the inhibition case, that is $\boldsymbol{\Theta} = \mathbb{R}^2 \times \mathbb{R}_+^3$ and $D_0^{m_1,m_2} = 0$. The following proofs could easily be extended to the hard-core case and to the multi-Strauss marked point process (see *e.g.* Billiot et al. (2008)). For any vector $\mathbf{z}$ of length 5, we sometimes reparameterize it similarly as the parameter vector, that is $\mathbf{z} = (z_1^1, z_1^2, z_2^{1,1}, z_2^{1,2}, z_2^{2,2})^T$.

B.1.1 Proof that $\underline{\boldsymbol{\Sigma}}_1(\theta^\star)$ is positive-definite for the two-type Strauss model

From Proposition 7, proving that $\underline{\boldsymbol{\Sigma}}_1(\theta^\star)$ is positive-definite in Framework 1 reduces to check Assumption **[PD]** with $h = e^V$ and

(*i*) $\mathbf{Y}_\Lambda(\varphi;\theta^\star) = I_\Lambda(\varphi;e^V,\theta^\star)$,

(*ii*) $\mathbf{Y}_\Lambda(\varphi;\theta^\star) = R_{\infty,\Lambda}(\varphi;e^V,\theta^\star)$.

(i) is ensured by Proposition 16.

(ii) We fix $\overline{\delta} = D$ and $B = \emptyset$ in **[PD]**. Let $\overline{\Omega} := \overline{\Omega}_{\emptyset}$. Without loss of generality, one may assume that $\theta_2^{\star 1,1} > 0$. Let us define for $n \geq 1$

$$A_{n,-}(\eta) = \Big\{\varphi \in \overline{\Omega} : \varphi(\Delta_0(\overline{\delta}) \times \{1\}) = 2n, \varphi(\Delta_0(\overline{\delta}) \times \{2\}) = 0, \\ \varphi\left(\mathcal{B}\left((0,0), \frac{\eta}{4}\right)\right) = n, \varphi\left(\mathcal{B}\left((D^{1,1} - \frac{\eta}{2}, 0), \frac{\eta}{4}\right)\right) = n\Big\},$$

$$A_{n,+}(\eta) = \Big\{\varphi \in \overline{\Omega} : \varphi(\Delta_0(\overline{\delta}) \times \{1\}) = 2n, \varphi(\Delta_0(\overline{\delta}) \times \{2\}) = 0, \\ \varphi\left(\mathcal{B}\left((0,0), \frac{\eta}{4}\right)\right) = n, \varphi\left(\mathcal{B}\left((D^{1,1} + \frac{\eta}{2}, 0), \frac{\eta}{4}\right)\right) = n\Big\}.$$

Let $\varphi_{n,-} \in A_{n,-}$ and $\varphi_{n,+} \in A_{n,+}$. Then for η small enough

$$I_{\overline{\Lambda}}(\varphi_{n,\bullet}; e^V, \theta^\star) = |\overline{\Lambda}| - \begin{cases} 2ne^{\theta_1^{\star 1} + (2n-1)\theta_2^{\star 1,1}} & \text{if } \bullet = -, \\ 2ne^{\theta_1^{\star 1} + (n-1)\theta_2^{\star 1,1}} & \text{if } \bullet = +. \end{cases}$$

$$\left(\mathbf{LPL}_{\overline{\Lambda}}^{(1)}(\varphi_{n,\bullet}; \theta^\star)\right)_1^{m'} = \int_{\overline{\Lambda} \times \mathbb{M}} v_1^{m'}(x^m|\varphi_{n,\bullet}) e^{-V(x^m|\varphi_{n,\bullet};\theta^\star)} \mu(dx^m) - \begin{cases} 2n & \text{if } m' = 1, \\ 0 & \text{if } m' = 2. \end{cases}$$

$$\left(\mathbf{LPL}_{\overline{\Lambda}}^{(1)}(\varphi_{n,-}; \theta^\star)\right)_2^{m_1',m_2'} = \int_{\overline{\Lambda} \times \mathbb{M}} v_2^{m_1,m_2}(x^m|\varphi_{n,-}) e^{-V(x^m|\varphi_{n,-};\theta^\star)} \mu(dx^m) \\ - \begin{cases} 2n(2n-1) & \text{if } m_1 = m_2 = 1, \\ 0 & \text{otherwise.} \end{cases}$$

$$\left(\mathbf{LPL}_{\overline{\Lambda}}^{(1)}(\varphi_{n,+}; \theta^\star)\right)_2^{m_1,m_2} = \int_{\overline{\Lambda} \times \mathbb{M}} v_2^{m_1,m_2}(x^m|\varphi_{n,+}) e^{-V(x^m|\varphi_{n,+};\theta^\star)} \mu(dx^m) \\ - \begin{cases} n(n-1) & \text{if } m_1 = m_2 = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Now,

$$\begin{aligned} &\Delta R_{\infty,\overline{\Lambda}}(\varphi_{n,-}, \varphi_{n,+}) := R_{\infty,\overline{\Lambda}}(\varphi_{n,-}; e^V, \theta^\star) - R_{\infty,\overline{\Lambda}}(\varphi_{n,+}; e^V, \theta^\star) \\ &= 2n\left(e^{\theta_1^{\star 1} + (n-1)\theta_2^{\star 1,1}} - e^{\theta_1^{\star 1} + (2n-1)\theta_2^{\star 1,1}}\right) + \left(\mathbf{W}(e^V, \theta^\star)\right)_2^{1,1}(2n(2n-1) - n(n-1)) \\ &\quad + f(\varphi_{n,-}, \varphi_{n,+}, \mathbf{W}, \eta) \\ &= \underbrace{2ne^{\theta_1^{\star 1} + (n-1)\theta_2^{\star 1,1}}(1 - e^{n\theta_2^{\star 1,1}}) + n(3n-1)\left(\mathbf{W}(e^V, \theta^\star)\right)_2^{1,1}}_{:=x_n} + f(\varphi_{n,-}, \varphi_{n,+}, \mathbf{W}, \eta). \end{aligned}$$

Fix $\varepsilon > 0$, there exists $n_0 \geq 1$ such that for all $n \geq n_0$, $x_n < -\varepsilon$. Now by a continuity argument, there exists $\eta_0 = \eta_0(n_0)$ such that for all $\eta \leq \eta_0(n_0)$, $|f(\varphi_{n_0,-}, \varphi_{n_0,+}, \mathbf{W}, \eta)| \leq \varepsilon/2$. Therefore by assuming that $\Delta R_{\infty,\overline{\Lambda}}(\varphi_{n_0,-}, \varphi_{n_0,+}) = 0$, we obtain for $\eta \leq \eta_0$

$$0 = |\Delta R_{\infty,\overline{\Lambda}}(\varphi_{n_0,-}, \varphi_{n_0,+})| \geq |x_{n_0}| - |f(\varphi_{n_0,-}, \varphi_{n_0,+}, \mathbf{W}, \eta)| \geq \varepsilon/2 > 0$$

which leads to a contradiction and proves **[PD]**.

B.1.2 Proof that $\mathbf{\Sigma}_2(\theta^\star)$ is positive-definite for the two-type Strauss model

From Proposition 7, proving that $\mathbf{\Sigma}_2(\theta^\star)$ is positive-definite in Framework 2 reduces to check Assumption **[PD]** with

$$\mathbf{Y}_\Lambda(\varphi; \theta^\star) = \mathbf{R}_{\infty,\Lambda}(\varphi; \mathbf{h}; \theta^\star),$$

where, for all $j = 1, \dots, s$, $(\mathbf{R}_{\infty,\Lambda}(\varphi; \mathbf{h}; \theta^\star))_j = R_{\infty,\Lambda}(\varphi; h_j, \theta^\star)$, h_j is the test function given by $h_j(x^m, \varphi; \theta) = \mathbf{1}_{[0,r_j]}(d(x^m, \varphi))e^{V(x^m|\varphi;\theta)}$. We fix as before $\overline{\delta} = D$ and $B = \emptyset$ in **[PD]**.

Let $0 < r_1 < \dots < r_s < +\infty$. Let us also assume that $r_i \neq D$ for $i = 1, \dots, s$ and define

$$A^{1,1}_{i,-}(\eta) = \left\{\varphi \in \overline{\Omega} : \varphi(\Delta_0(\overline{D})) = 2, \varphi\left(\mathcal{B}\left((0,0), \frac{\eta}{4}\right) \times \{1\}\right) = 1, \varphi\left(\mathcal{B}\left((r_i - \frac{\eta}{2}, 0), \frac{\eta}{4}\right) \times \{1\}\right) = 1\right\},$$
$$A^{1,1}_{i,+}(\eta) = \left\{\varphi \in \overline{\Omega} : \varphi(\Delta_0(\overline{D})) = 2, \varphi\left(\mathcal{B}\left((0,0), \frac{\eta}{4}\right) \times \{1\}\right) = 1, \varphi\left(\mathcal{B}\left((r_i + \frac{\eta}{2}, 0), \frac{\eta}{4}\right) \times \{1\}\right) = 1\right\}.$$

Let $\varphi_{i,\bullet} \in A^{1,1}_{i,\bullet}(\eta)$ for $\bullet = -, +$ and i$= 1, \dots, s$. Let κ_i the constant given by

$$\kappa_i = \begin{cases} 2e^{\theta_1^{\star 1} + \theta_2^{\star 1,1}} & \text{if } r_i < D, \\ 2e^{\theta_1^{\star 1}} & \text{otherwise.} \end{cases}$$

Then for $i, j = 1, \dots, s$ and for η small enough

$$I_{\overline{\Lambda}}(\varphi_{i,-}; h_j, \theta^\star) = \int_{\overline{\Lambda} \times \mathbb{M}} h_j(x^m, \varphi_{i,-})e^{-V(x^m|\varphi_{i,-};\theta^\star)}\mu(dx^m) - \begin{cases} \kappa_i & \text{if } i \leq j, \\ 0 & \text{otherwise.} \end{cases}$$
$$I_{\overline{\Lambda}}(\varphi_{i,+}; h_j, \theta^\star) = \int_{\overline{\Lambda} \times \mathbb{M}} h_j(x^m, \varphi_{i,+})e^{-V(x^m|\varphi_{i,+};\theta^\star)}\mu(dx^m) - \begin{cases} \kappa_i & \text{if } i < j, \\ 0 & \text{otherwise.} \end{cases}$$

On the other hand

$$\left(\mathbf{LPL}^{(1)}_{\overline{\Lambda}}(\varphi_{i,\bullet}; \theta^\star)\right)^{m'}_1 = \int_{\overline{\Lambda} \times \mathbb{M}} v_1^{m'}(x^m|\varphi_{i,\bullet} e^{-V(x^m|\varphi_{i,\bullet};\theta^\star)}\mu(dx^m) - \begin{cases} 2 & \text{if } m' = 1, \\ 0 & \text{if } m' = 2. \end{cases}$$
$$\left(\mathbf{LPL}^{(1)}_{\overline{\Lambda}}(\varphi_{i,\bullet}; \theta^\star)\right)^{m_1,m_2}_2 = \int_{\overline{\Lambda} \times \mathbb{M}} v_2^{m_1,m_2}(x^m|\varphi_{i,\bullet} e^{-V(x^m|\varphi_{i,\bullet};\theta^\star)}\mu(dx^m) - \begin{cases} 2 & \text{if } m_1 = m_2 = 1 \\ & \text{and } r_i < D \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathbf{x} \in \mathbb{R}^s \setminus \{0\}$, then from previous computations

$$\mathbf{x}^T\left(\mathbf{R}_{\infty,\overline{\Lambda}}(\varphi_{i,+}; \mathbf{h}, \theta^\star) - \mathbf{R}_{\infty,\overline{\Lambda}}(\varphi_{i,-}; \mathbf{h}, \theta^\star)\right) = 2\kappa_i x_i + f(\mathbf{x}, \varphi_{i,+}, \varphi_{i,-}, \mathbf{h}). \tag{56}$$

By using a continuity argument, one may prove that for every $\varepsilon > 0$ there exists $\eta > 0$ such that $|f((\mathbf{x}, \varphi_{i,+}, \varphi_{i,-}, \mathbf{h})| \leq \varepsilon$. Therefore, assuming that the l.h.s. of (57) equals 0 leads to $x_i = 0$ for $i = 1, \dots, s$.

B.2 Area-interaction point process

We fix for simplicity $d = 2$, though the proofs may be extended easily to higher dimensions.

B.2.1 Proof that $\underline{\boldsymbol{\Sigma}}_1(\theta^\star)$ is positive-definite for the area-interaction model

From Proposition 7, the proof reduces to check Assumption **[PD]** with $h = e^V$, $\overline{\delta} = D$, $B = \emptyset$ and

(*i*) $\mathbf{Y}_\Lambda(\varphi; \theta^\star) = I_\Lambda(\varphi; e^V, \theta^\star)$,

(*ii*) $\mathbf{Y}_\Lambda(\varphi; \theta^\star) = R_{\infty,\Lambda}(\varphi; e^V, \theta^\star)$.

Again (*i*) is ensured by Proposition 16 since this model satisfies **[Exp]**.

(*ii*) Let us consider for some $\eta, \omega > 0$ the two following events:

$$\begin{aligned} A_1(\eta, \omega) &:= \left\{\varphi \in \overline{\Omega} : \varphi(\Delta_0(\overline{\delta})) = 2, \varphi(\mathcal{B}((0,0), \eta)) = 1, \varphi(\mathcal{B}((0,\omega), \eta)) = 1\right\} \\ A_2(\eta, \omega) &:= \left\{\varphi \in \overline{\Omega} : \varphi(\Delta_0(\overline{\delta})) = 3, \varphi(\mathcal{B}((0,0), \eta)) = 1, \varphi(\mathcal{B}((0,\omega), \eta)) = 2\right\} \end{aligned}$$

Fix η, ω, let $\varphi_j \in A_j(\eta,\omega)$ and denote by $\widetilde{e^V}(\varphi) := \sum_{x\in\varphi_{\overline{\Lambda}}} e^{V(x|\varphi\setminus x;\theta^\star)}$

$$I_{\overline{\Lambda}}(\varphi_j; e^V, \theta^\star) \quad = \quad |\overline{\Lambda}| - \widetilde{e^V}(\varphi_j).$$

When $\eta \to 0$,

$$\widetilde{e^V}(\varphi_1) \to 2e^{\theta_1^\star + \theta_2^\star g(\omega)} \quad \text{and} \quad \widetilde{e^V}(\varphi_2) \to 2e^{\theta_1^\star} + e^{\theta_1^\star+\theta_2^\star g(\omega)}$$

where $g(\omega) := |\mathcal{B}((0,0),R) \cup \mathcal{B}((0,\omega),R)| - |\mathcal{B}((0,0),R)|$. Moreover, by denoting $\widetilde{v}_2(\varphi) = \sum_{x\in\varphi} v_2(x|\varphi\setminus x)$

$$\begin{aligned}
\left(\mathbf{LPL}^{(1)}_{\overline{\Lambda}}(\varphi_j;\theta^\star)\right)_1 &= \int_{\overline{\Lambda}} e^{-V(x|\varphi_j;\theta^\star)}dx - \left\{ \begin{array}{ll} 2 & \text{if } j=1 \\ 3 & \text{if } j=2 \end{array}\right. \\
\left(\mathbf{LPL}^{(1)}_{\overline{\Lambda}}(\varphi_j;\theta^\star)\right)_2 &= \int_{\overline{\Lambda}} v_2(x|\varphi_j) e^{-V(x|\varphi_j;\theta^\star)}dx - \widetilde{v}_2(\varphi_j).
\end{aligned}$$

Again, when $\eta \to 0$, one may note that for $k=1,2$

$$\int_{\overline{\Lambda}} v_k(x|\varphi_1)e^{-V(x|\varphi_1;\theta^\star)}dx - \int_{\overline{\Lambda}} v_k(x|\varphi_2)e^{-V(x|\varphi_2;\theta^\star)}dx \to 0$$

and

$$\widetilde{v}_2(\varphi_1) \to 2g(\omega) \quad \text{and} \quad \widetilde{v}_2(\varphi_2) \to g(\omega)$$

These computations lead to

$$\begin{aligned}
R_{\infty,\overline{\Lambda}}(\varphi_1; e^V,\theta^\star) - R_{\infty,\overline{\Lambda}}(\varphi_2; e^V,\theta^\star) &= 2e^{\theta_1^\star} - e^{\theta_1^\star+\theta_2^\star g(\omega)} - \left(\mathbf{W}(e^V,\theta^\star)\right)_1 + g(\omega)\left(\mathbf{W}(e^V,\theta^\star)\right)_2 \\
&\quad + f(\varphi_1,\varphi_2,\mathbf{W}),
\end{aligned}$$

where the function f is such that for all $\varepsilon > 0$, there exists η small enough such that $|f(\varphi_1,\varphi_2,\mathbf{W})| \le \varepsilon$. Let $\varphi_j \in A_j(\eta, 0)$, then, since $g(0)=0$, assuming that the l.h.s. of the previous equation equals 0 leads to $\left(\mathbf{W}(e^V,\theta^\star)\right)_1 = e^{\theta_1^\star}$. Now, let $\omega > 0$ and again assume that $R_{\infty,\overline{\Lambda}}(\varphi_1; e^V,\theta^\star) = R_{\infty,\overline{\Lambda}}(\varphi_2; e^V,\theta^\star)$, we therefore obtain (by the continuity argument)

$$\left(\mathbf{W}(e^V,\theta^\star)\right)_2 = \frac{e^{\theta_1^\star+\theta_2^\star g(\omega)} - e^{\theta_1^\star}}{g(\omega)}.$$

But $\left(\mathbf{W}(e^V,\theta^\star)\right)_2$ is a constant and so cannot depend on ω. Therefore one of the assumptions made before is untrue, which proves **[PD]**.

B.2.2 Proof that $\underline{\Sigma}_2(\theta^\star)$ is positive-definite for the area-interaction model

From Proposition 7, it suffices to check Assumption **[PD]** with

$$\mathbf{Y}_\Lambda(\varphi;\theta^\star) = \mathbf{R}_{\infty,\Lambda}(\varphi;\mathbf{h};\theta^\star),$$

where, for all $j=1,\dots,s$, $(\mathbf{R}_{\infty,\Lambda}(\varphi;\mathbf{h};\theta^\star))_j = R_{\infty,\Lambda}(\varphi;h_j,\theta^\star)$, h_j is the test function given by $h_j(x^m,\varphi;\theta) = \mathbf{1}_{[0,r_j]}(d(x^m,\varphi))e^{V(x^m|\varphi;\theta)}$, and where, again, we choose $\overline{\delta} = D$ and $B = \emptyset$.

The proof is quite similar to the one proposed for the 2-type marked Strauss point process (see B.2.2). Let $0 < r_1 < \ldots < r_s < +\infty$. Let us also assume that $r_i \neq D$ for $i=1,\dots,s$

$$\begin{aligned}
A_{i,-}(\eta) &= \left\{\varphi \in \overline{\Omega} : \varphi(\Delta_0(\overline{D})) = 2, \varphi\left(\mathcal{B}\left((0,0),\frac{\eta}{4}\right)\right) = 1, \varphi\left(\mathcal{B}\left((r_i - \frac{\eta}{2},0),\frac{\eta}{4}\right)\right) = 1\right\}, \\
A_{i,+}(\eta) &= \left\{\varphi \in \overline{\Omega} : \varphi(\Delta_0(\overline{D})) = 2, \varphi\left(\mathcal{B}\left((0,0),\frac{\eta}{4}\right)\right) = 1, \varphi\left(\mathcal{B}\left((r_i + \frac{\eta}{2},0),\frac{\eta}{4}\right)\right) = 1\right\}.
\end{aligned}$$

Let $i,j \in \{1,\ldots,s\}$ and $k \in \{1,2\}$, let $\varphi_{i,-} \in A_{i,-}$ and $\varphi_{i,+} \in A_{i,+}$, then

$$\begin{aligned} I_{\overline{\Lambda}}(\varphi_{i,-};h_j,\theta^\star) &= \int_{\overline{\Lambda}} h_j(x,\varphi_{i,-};\theta^\star)e^{-V(x|\varphi_{i,-};\theta^\star)}dx - \begin{cases} \widetilde{e^V}(\varphi_{i,-}) & \text{if } i \le j \\ 0 & \text{otherwise.} \end{cases} \\ I_{\overline{\Lambda}}(\varphi_{i,+};h_j,\theta^\star) &= \int_{\overline{\Lambda}} h_j(x,\varphi_{i,+};\theta^\star)e^{-V(x|\varphi_{i,+};\theta^\star)}dx - \begin{cases} \widetilde{e^V}(\varphi_{i,+}) & \text{if } i < j \\ 0 & \text{otherwise.} \end{cases} \\ \left(\mathbf{LPL}^{(1)}_{\overline{\Lambda}}(\varphi_{i;\bullet};\theta^\star)\right)_k &= \int_{\overline{\Lambda}} v_k(x|\varphi_{i,\bullet})e^{-V(x|\varphi_{i,\bullet};\theta^\star)}dx - \sum_{x\in\varphi_{i,\bullet}} v_k(x|\varphi_{i,\bullet}\setminus x), \end{aligned}$$

for $\bullet = -,+$. It is expected that for small η, $\left(\mathbf{LPL}^{(1)}_{\overline{\Lambda}}(\varphi_{i,-};\theta^\star)\right)_k \simeq \left(\mathbf{LPL}^{(1)}_{\overline{\Lambda}}(\varphi_{i,+};\theta^\star)\right)_k$ and $\widetilde{e^V}(\varphi_{i,-}) \simeq \widetilde{e^V}(\varphi_{i,+}) \simeq \kappa_i := 2e^{\theta_1^\star+\theta_2^\star|\mathcal{B}(0,R)\cup\mathcal{B}(r_i,R)|}$. Let $\mathbf{x} \in \mathbb{R}^s \setminus \{0\}$, then from previous computations

$$\mathbf{x}^T\left(\mathbf{R}_{\infty,\overline{\Lambda}}(\varphi_{i,+};\mathbf{h},\theta^\star) - \mathbf{R}_{\infty,\overline{\Lambda}}(\varphi_{i,-};\mathbf{h},\theta^\star)\right) = 2\kappa_i x_i + f(\mathbf{x},\varphi_{i,+},\varphi_{i,-},\mathbf{h}) \tag{57}$$

where for every $\varepsilon > 0$ there exists $\eta > 0$ such that $|f((\mathbf{x},\varphi_{i,+},\varphi_{i,-},\mathbf{h})| \le \varepsilon$. Therefore, assuming that the l.h.s. of (57) equals 0 leads to $x_i = 0$ for $i = 1,\ldots,s$.

References

P.K. Andersen, Ø Borgan, R.D. Gill, and N. Keiding. *Statistical models based on counting processes.* Springer-Verlag, New York, 1993.

A. Atkinson. *Plots, Transformations and Regression.* Oxford Statistical Sciences. Oxford University Press/ Clarendon, 1985.

A. Baddeley and R. Turner. Practical maximum pseudolikelihood for spatial point patterns (with discussion). *Australian and New Zealand Journal of Statistics*, 42:283–322, 2000.

A. Baddeley and R. Turner. Modelling spatial point patterns in R. *Journal of Statistical Software*, 12(6):1–42, 2005.

A. Baddeley, R. Turner, J. Møller, and M. Hazelton. Residual analysis for spatial point processes. *Journal of the Royal Statistical Society (series B)*, 67:1–35, 2005.

A. Baddeley, J. Møller, and A.G. Pakes. Properties of residuals for spatial point processes. *Annals of the Institute of Statistical Mathematics*, 60(3):627–649, 2008.

E. Bertin, J.-M. Billiot, and R. Drouilhet. Existence of "Nearest-Neighbour" Gibbs Point Models. *Ann. Appl. Probab.*, 31:895–909, 1999.

J.-M. Billiot, J.-F. Coeurjolly, and R. Drouilhet. Maximum pseudolikelihood estimator for exponential family models of marked Gibbs point processes. *Electronic Journal of Statistics*, 2: 234–264, 2008.

E. Bolthausen. On the central limit theorem for stationary mixing random fields. *The Annals of Probability*, 10(4):1047–1050, 1982.

J.-F. Coeurjolly and R. Drouilhet. Maximum pseudo-likelihood estimator for general marked Gibbs point processes and applications to the lennard-jones model. *submitted*, submitted (http://hal.archives-ouvertes.fr/hal-00442750/en/), 2009.

F. Comets and M. Janzura. A central limit theorem for conditionally centred random fields with an application to Markov fields. *Journal of Applied Probability*, 35(3):608–621, 1998.

D. Daley and D. Vere-Jones. *An introduction to the Theory of Point Processes*. Springer Verlag, New York, 1988.

D. Dereudre. Gibbs Delaunay tessellations with geometric hardcore condition. *J. Stat. Phys.*, 121 (3-4):511–515, 2005.

D. Dereudre and F. Lavancier. Campbell equilibrium equation and pseudo-likelihood estimation for non-hereditary Gibbs point processes. *Bernoulli*, 15(4):1368–1396, 2009.

D. Dereudre and F. Lavancier. Practical simulation and estimation for Gibbs Delaunay-voronoï tessellations with geometric hardcore interaction. *submitted (http://hal.archives-ouvertes.fr/hal-00402029/en/)*, 2010.

D. Dereudre, R. Drouilhet, and H.O. Georgii. Existence of Gibbsian point processes with geometry-dependent interactions. *submitted*, 2010.

P.J. Diggle. *Statistical analysis of spatial point patterns*. Academic Press, London, 2003.

T.R. Fleming and D.P. Harrington. *Counting processes and survival analysis*. Wiley, New York, 1991.

H.O. Georgii. Canonical and grand canonical Gibbs states for continuum systems. *Communications in Mathematical Physics*, 48(1):31–51, 1976.

J. Illian, A. Penttinen H., and Stoyan. *Statistical analysis and modelling of spatial point patterns*. Wiley-Interscience, 2008.

J.L. Jensen and H.R. Künsch. On asymptotic normality of pseudo likelihood estimates of pairwise interaction processes. *Ann. Inst. Statist. Math.*, 46:475–486, 1994.

J.L. Jensen and J. Møller. Pseudolikelihood for exponential family models of spatial point processes. *Ann. Appl. Probab.*, 1:445–461, 1991.

J. Møller. Parametric methods for spatial point processes. Technical Report Research Report R-2008-04, Department of Mathematical Sciences, Aalborg University, 2008.

J. Møller and R. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton, 2003.

X. Nguyen and H. Zessin. Integral and differential characterizations Gibbs processes. *Mathematische Nachrichten*, 88(1):105–115, 1979a.

X.X. Nguyen and H. Zessin. Ergodic theorems for Spatial Process. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 48:133–158, 1979b.

C.J. Preston. *Random fields*. Springer Verlag, 1976.

D. Ruelle. *Statistical Mechanics*. Benjamin, New York-Amsterdam, 1969.

D. Stoyan and P. Grabarnik. Second-order characteristics for stochastic structures connected with Gibbs point processes. *Mathematische Nachrichten*, 151:95–100, 1991.

D. Stoyan, W.S. Kendall, J. Mecke, and L. Ruschendorf. *Stochastic geometry and its applications*. John Wiley and Sons, Chichester, 1987.